

29 July 1982

# Watch the 17 per cent solution

*Spokesmen for United States science have expressed delight with the recent increase of research spending by industry. But their optimism is ill-founded and may be premature.*

Has United States industry changed its attitude towards research? After years of only modest increases of spending on research and development and often of self-confessed neglect, corporations have now increased their spending mightily. A much discussed and recently published McGraw-Hill survey has shown that corporate spending will increase in 1982 by 17 per cent. If inflation this year remains at 6.5 per cent, the real increase will be more than 10 per cent — and this on the heels of an increase of 16 per cent in 1981. So some spokesmen for US science, surveying the mixed outlook for government support that will persist for the foreseeable future (see *Nature* 8 July, p.112) have greeted the new industry money with glee. For example, Dr Frank Press, president of the National Academy of Sciences, calls them signs of “a major change in attitude”. McGraw-Hill, in releasing its survey, called it evidence of “an R&D renaissance”. But it may be premature to pat industrial managers on the back for new-found wisdom.

Three caveats are important. The first is uncertainty about the working of the Economic Recovery Tax Act, signed by President Reagan in April 1981 and intended, among other things, to spur investment in research. No definitive analysis has been made of the act's specific impact on corporations, although industry leaders acknowledge that it has been considerable. If, for example, a corporation adds a new wing to a plant that includes some experimental production techniques, it may count as a research and development item and qualify for special tax treatment. Corporation accountants would be crazy not to take every advantage they can of the act, which has probably swollen many corporate statements of research spending without causing research itself — let alone long-term research — to grow. So much of the increase may be a paper benefit.

A second caveat is the dog that didn't bark. What has happened to the large fraction of industrial research and development that has to go for “defensive” research to get new products or processes through the mass of federal regulation? For years, in the 1970s, industry spokesmen used to complain in no uncertain terms that, if only government would lift the regulatory burden, industry could stop spending lopsided amounts on this short-term work and invest in the long-term high-risk innovative research that could keep its long-term trade position secure. Now, quite suddenly, industrial managers have stopped complaining. What is really going on? The Reagan Administration has lifted many regulatory burdens, to be sure, but just as many remain, particularly affecting the chemical and pharmaceutical industries. Have companies decreased the amount of “defensive” research they are doing and, as they promised, switched to more long-term projects? That seems unlikely, especially as some industry researchers have been saying privately that they are now under more pressure than ever to produce quick results. Most probably, industry's habits have not changed much, that just as much (or more) “defensive” research is being done and that industry spokesmen are quieter because an administration sympathetic to their complaints is in office. The argument that federal regulation has prevented companies from doing long-term research may have been a hoax in the first place.

A third and final ground for caution is the Administration's defence build-up, which is sprinkling money into the electronics, information and aerospace industries like spring rain. Predictably, in the McGraw-Hill survey, major defence

contractors such as IBM, Boeing, United Technologies and Cray Research lead the pack. Robert Reich, of Harvard University's John F. Kennedy School of Government, has eloquently pointed out how the Department of Defense acts as the American equivalent of Japan's Ministry of International Trade and Industry (MITI) by encouraging the development of key technologies such as lasers, integrated circuits and space gadgets. But whereas MITI assists these technologies in civilian markets directly, the Pentagon does so only indirectly and over the longer term. But, in the short run, massive Pentagon research spending distracts industry's attention from civilian products.

So the significance of the 17 per cent “solution” now being hailed is really unknown. Trends in United States industrial research bear serious study. Indeed, in view of the country's high technology trade balance, a close sceptical look is urgently necessary, for changes must be made now to improve industry's performance in the future. It is wrong to assume — as Americans often do — that just because money is being spent on a problem it is being solved.

Two structural problems having nothing to do with money should be looked at closely. One is the unsolved problem of US industry's relations with government in the development of high technology industries and products that could affect future trade. As Reich says, countries such as Japan that compete successfully in international markets have smooth well-understood relations between government and industry. The ball is passed between these sectors smoothly, as with a well-coached basket-ball team. In spite of the Reagan Administration's pro-industry stance, its easing of regulations and bowing to many of industry's special interests, the underlying stand-off between the private and public sectors in the United States remains. As a simple example, what will happen to solar energy technology in the United States, now that the Reagan Administration has dropped the government's contribution? Because of poor industry/government coordination in this field, earlier investment will be lost. Will the Japanese win that one too?

A second problem is internal to the companies themselves. Although many large corporations spend money on research and development, their top management may pay it little heed; research managers are told to mind their own business, which is “merely technical”, rather than to suggest changes that could change the way the corporation does business. General Motors (GM), for example, emerged as the top spender in the McGraw-Hill survey: it will spend an estimated \$2,200 million in 1982 (out of industry's total spending of \$59,700 million). But at GM “to be assigned to research means you're dead”. Clearly relations between top management and research people must be improved. How many chief executive officers, for example, have lately come from research rather than the law, accountancy or advertising?

The debate over industrial research raises questions, finally, about the much-touted liaison between US industry and universities. So far, it has the air of a flirtation rather than a stable marriage. If corporations *have* changed their attitudes, and are more seriously investing in long-term undirected research, they should see how much they have in common with the university research scene. Then the partnership can be stable, and not just kiss-and-tell. Spokesmen of science should be asking, have industrial managers become suddenly and miraculously wiser?



# Test-bans up and down

*The United States wants two test-ban treaties reopened. Why not just ratify them?*

Is the United States backsliding with its request to the Soviet Union that two bilateral treaties on nuclear weapons tests should be reopened? The measure of agreement that has been accumulated since the 1963 treaty banning tests in the atmosphere is one of the few achievements in three frustrating decades of arms control. Although the strategic arms reduction talks (Start) begun in Geneva suggest that the Reagan Administration has taken a step forward on arms control, it would rightly be counted a setback if the test-ban treaties were to be thrown back into the melting pot. So what is the Administration doing? Is the reopening a prelude to making the treaties acceptable to the Senate? Or will the Administration wash its hands of them?

First, it is not intended that the 1963 treaty should be reopened. That could not in any case be done without the agreement of the other signatories. Second, the Administration seems to be concerned only with the two related bilateral agreements between itself and the Soviet Union — the agreement signed in 1974 by which each superpower agreed not to test weapons yielding more than the equivalent of 150,000 tonnes of conventional explosive, and that signed in 1976 on peaceful nuclear explosions. Neither of these treaties has been formally ratified, but both parties to the treaties seem ever since to have kept the rules. The United States now wants to improve the provisions for verification on the grounds that the technology has improved in the past few years — and that the provisions of the treaties as they stand are too loose to be convincing to the US Senate.

That is too cheerful a reading of what has happened. The treaties now up again for grabs are very different from each other. The threshold treaty stipulates that verification should depend on "national technical means" — remote seismographs and the like — although a protocol to the treaty (which would have the force of law if the treaty were ever ratified) requires that each side should provide the other with data to help with the calibration of these external measurements. Washington's chief concern seems to be that many Soviet tests of weapons underground have been uncomfortably close to the limit of 150,000 tonnes, at least while errors of measurement are as great as at present. Obviously the treaty would be improved if each side would agree to the siting of automatic monitoring stations at each other's test sites (which, under the treaty, must be specified geographically in advance). If that is what the United States plans to ask for, well and good.

The treaty on peaceful nuclear explosions is a different kettle of fish. Its chief objective is to ensure that peaceful explosions are not used as a means of circumventing the threshold treaty. Among international agreements of this kind, however, this treaty includes some of the most stringent means of verification ever to be agreed. Those wishing to carry out explosions are required not merely to give advance notice of their intentions (including the time of detonation to within a second) but, for explosions yielding between 100 and 150 kilotonnes, to admit inspectors from the other side. Since only the Soviet Union appears to be interested in such explosions (there appear to have been five in 1981), it has seemed on many occasions in the past few years that the United States could win a unilateral benefit by ratifying the treaty and bringing the disclosure provisions into force. If the United States's objective is merely to extend these provisions to smaller explosions, nobody will complain — although there is a risk that to ask for too much will make the Soviet Union retreat from the concessions already made.

The comprehensive test-ban treaty is something else again. When negotiations ground to a halt two years ago, many in President Carter's Administration were fond of saying that a treaty could be completed within two months, given the "political will". This opinion was always too optimistic, for neither President Carter nor his advisers had accurately calculated the chances of persuading the military to forgo tests of any kind for the indefinite future. (Something can be done to improve the

design of nuclear weapons tests by tests of conventional explosives, sub-critical experiments and mere calculation, but the process is uncertain.) The Reagan Administration has now said that it will not pursue the comprehensive treaty. That is not surprising, for it recognizes what is feasible in Washington.

That does not mean that the comprehensive test-ban is dead. The technical study of the international monitoring of a comprehensive test-ban treaty set up earlier this year by the Geneva Committee on Disarmament has terms of reference that will allow it to make the best possible case for supposing that a comprehensive test-ban could be monitored effectively. Whether France and China, the two nuclear powers still outside the 1963 test-ban treaty, will attend the technical session beginning on 3 August remains to be seen. Whatever happens, there is now a good chance that the Geneva committee will be able to back a case that, perhaps with the help of automatic monitoring stations (part of the Carter draft), uncertainty about verification is no longer a valid reason for not having a comprehensive test-ban. The timetable is about right to cause the maximum embarrassment to the nuclear powers at the next review conference of the Non-Proliferation Treaty in 1985.

## Britain's test-tube babies

*The new British committee on in vitro fertilization should not make heavy weather of its task.*

In one respect at least, British technology has established a commanding lead: children conceived by the technology of *in vitro* fertilization are being born at a greater rate than anywhere else (see *Nature* 293, 253; 1981). Predictably, the British government has therefore set up a committee (under Dr Mary Warnock, a lively soul) to consider the implications, ethical and otherwise, and to recommend what should be done. This is what the committee's report should say.

- Devised as a means of treating certain kinds of female sterility, *in vitro* fertilization works and produces predominantly healthy children, but naturally more data on the last point would be welcome. For the small proportion of women thus handicapped, the treatment is demonstrably beneficial and should be provided as if it were a medical routine.

- As a technique for helping sterile women to produce children, *in vitro* fertilization is less open to theoretical objections than other techniques now in use. Thus circumventing male sterility by means of AID (artificial insemination by a donor) is widely practised and is to some people objectionable because the children's DNA is not some reshuffling (by the rules of genetic recombination) of the parental DNA, which is an old-fashioned grumble. The practice of AID nevertheless lends itself to such abuses as those of genetically-attested sperm banks, which should be registered and then regulated.

- Although *in vitro* fertilization could in principle be used to implant surrogate mothers with embryos derived from two other people, thus creating an unseemly trade, it must surely be significant that this has not happened when female sterility cannot be treated by this or other means.

- The use and disposal of surplus embryos produced by *in vitro* fertilization is a problem for many religious people, Roman Catholics in particular. If it were accepted that each viable fertilized ovum is human and alive, not to implant it is to sin. But this strict view cannot be made universally applicable in societies in which abortion is permitted. The more difficult question is whether viable human embryos may be used in embryological studies; the committee should say yes but should fix a time limit corresponding to the normal time of implantation for the independent growth in culture of human embryos, recommending that this time limit will be extended only when animal experiments show that embryologists have pointed questions to ask.

- Otherwise (the committee should say) *in vitro* fertilization should not be a public worry. Human cloning, when it becomes feasible, will be different.



# Key biotechnology patent delayed

## Cohen-Boyer patents could face challenge

Washington

The US Patent and Trademark Office suddenly announced on 30 June that it was delaying the issue of a second fundamental patent on genetic engineering to Stanford University and the University of California. The patent, which complements the Cohen-Boyer patent issued to the universities in December 1980, had already been assigned a patent number and was scheduled for issue on 13 July.

The announcement immediately sparked off speculation that a challenge to the patents was imminent. That speculation was fuelled by the revelation that an Exxon patent attorney, Albert Halluin, had discovered several potential defects in the original patent. Halluin's findings are to appear in a book called *The Patenting of Life Forms*, to be published by Cold Spring Harbor Laboratory on 15 August.

The most potentially damaging of Halluin's observations is that since the original patent was filed by Dr Stanley Cohen of Stanford and Dr Herbert Boyer of the University of California, San Francisco in November 1974, new findings have invalidated some of the information supplied in the patent specifications. The patent covers a key plasmid which is used as the vehicle for inserting new genes into the bacterium *Escherichia coli* and describes the method for producing this plasmid. But in 1977, Dr Cohen published a paper in *Journal of Bacteriology* that admits an error in the original procedure. A number of molecular geneticists have asserted that the error was substantial enough to make it impossible to duplicate Cohen's and Boyer's work from the patent description alone.

Halluin argues that deposition of the product in the American Type Culture Collection can get around such problems, but — and this is a second potential flaw in the patent — Cohen and Boyer did not deposit their plasmid until June 1981, more than six months after the patent was issued.

A third possible defect in the original patent concerns prior disclosure. A patent is not granted if information sufficient to duplicate the process is made public more than a year before application is made. Halluin points out that an article in *New Scientist* on 25 October 1973, more than a year before the application was filed, gives a detailed report of Cohen's and Boyer's work from what was supposed to have been an off-the-record Gordon Conference.

Halluin is chairman of the Chemical

Practice Committee of the American Patent Law Association, and the chapter he contributed to the Cold Spring Harbor book — of which only a very minor part deals with the Stanford patent — was apparently written in this capacity and not as an employee of Exxon Research and Engineering.

A spokesman for Exxon said it does not plan to challenge the validity of the Stanford patent and is not even officially studying it. He said the reason Exxon had not joined with 73 other companies in buying a one-year licence to the patent (at a price of \$10,000) was that it had no plans to use the process commercially.

In spite of the rumour of a challenge to the patents, none has been filed, according to Rene Tegtmeyer, the US assistant commissioner for patents. Tegtmeyer said he could not elaborate on the reasons for the delay in the second patent, but said "it happens a couple of hundred times a year" that a patent examiner requests a reexamination in the light of new information. An explanation will probably come in two or three weeks, when a patent



examiner is assigned to the reexamination. Normally, all such actions by the Patent Office are kept confidential. In this case, the applicant took the unusual step of opening its patent office file to public access.

Stanford's director of technology licensing, Niels Reimers, said that he was

## Chemical weapons denied funds

Washington

The House of Representatives voted last week by a wide margin against the production of binary chemical weapons.

The 251-159 vote came on an amendment to the defence authorization bill which deletes \$54 million that the Reagan Administration had requested to begin production. The United States has not manufactured chemical weapons since 1969, when President Nixon ordered that production should be halted. The strength of congressional opposition to breaking that 13-year moratorium was evident when 81 Republicans broke ranks with the Administration and voted to block the funds.

The Administration wanted the new weapons as a counter to what it sees as evidence of increased Soviet production of chemical warfare agents and an increased willingness to use them. Opposition was led by Representative Clement Zablocki (Democrat, Wisconsin) and Representative Ed Bethune (Republican, Arkansas), the sponsors of the amendment, who managed to use the Administration's own argument against it. "We have an opportunity to demonstrate that we are not like the Soviets", Bethune said. "It just doesn't make sense to throw away the one shred of evidence that Americans truly yearn for the day when arms will be reduced." Zablocki argued further that to proceed with production would divide the NATO alliance, as the Europeans — with the possible exception of France — are opposed to having new

chemical weapons on their soil. Opponents of production have argued that the new weapons are useless as a deterrent unless they are positioned in Europe.

The United States has a large stock of the older, unitary shells and bombs, which contain live nerve gas. Stocks are maintained in West Germany as well as the United States. The army says the binary weapons, which contain two relatively non-toxic gases that mix in flight to produce the nerve agent, are safer to store and to handle and are needed to replace deteriorating stocks of the older weapons.

The Senate, which in May approved the production of binary weapons by a close 49-45 vote, is likely to accede to the House view when the two chambers confer on a final authorization bill. The House action does not affect the \$705 million that the Administration is requesting for further research and development on binary weapons and for chemical defence.

Congress last year authorized construction of a facility in Arkansas to produce the new weapons, but stopped short of authorizing production. According to Bethune, the House action this year may have been more a response to public pressure — which has apparently been heightened by reports of Soviet use of chemical weapons in Afghanistan and Soviet complicity in the "yellow rain" episodes in South East Asia — than a response to logic and reason. As one staff member said, "no one wants to get up and speak in favour of nerve gas".

Stephen Budiansky



just as much in the dark about the delay as anybody. "In itself", however, "it's nothing momentous", he said. The issues raised by Halluin "have been looked at" and do not challenge the basic claim of Cohen and Boyer.

He also questioned the benefits to anybody of mounting a challenge. "Recall that the royalties are very low and no-one has contested that they [Cohen and Boyer] were the first ones to do this. And given the fact that everyone's getting a licence, I don't know why anybody would challenge it."

But royalties may not continue to be low. One industry observer pointed out that once commercial production begins — especially production of products with a high mark-up — royalties could be "substantial". And the incentive to mount a challenge would also grow. The licence contracts call for royalties of 1 per cent on net sales up to \$5 million,  $\frac{3}{4}$  per cent on the next \$5 million, and  $\frac{1}{2}$  per cent after that.

If a challenge is eventually filed, it could well be based on the issues that Halluin raised, and would probably take the form of a request for a reexamination of the first patent. Since December 1980, it has been possible to file such a request directly with the Patent Office, a much simpler procedure than mounting a lawsuit as was previously required. A recent Supreme Court ruling allows even a licensee to challenge a patent.

The original Stanford patent was a process claim, covering the production of transformants. The second patent, a product claim, covers the transformants themselves. In practice, however, it adds relatively little: the only additional protection it provides is in preventing a foreign company from producing transformants outside the patent and then selling them in the United States.

Stephen Budiansky

## Polish students

### Subjective change

There threatens to be a significant drop in the numbers of Polish students reading technical subjects in the next academic year, beginning in October. According to Warsaw radio, there has been a "qualitative transformation" of young people's attitudes to study. School-leavers, said the commentator, are thinking more frequently in "practical categories" and, in particular, of job prospects.

The swing away from technical studies does not, however, seem to be entirely spontaneous. At the end of June, the Minister of Science, Higher Education and Technology, Dr Benon Miskiewicz, said that the intake quota for the polytechnic universities had been cut this year due to a reduced demand for engineering personnel.

The minister's explanation, that the admission quotas have been deliberately

cut, seems the more likely. In Poland's current economic stagnation, many of the grandiose engineering projects of the Gierk regime have been suspended or, tacitly, cancelled. Moreover, since the alternative to higher education for male school-leavers is military service, it seems highly unlikely that many would deliberately opt for the latter.

Unlike many of its allies, Poland has never attempted to plan university admissions strictly in accordance with job prospects. The ministry has, however, the right to set limits to admissions in subjects that are over-subscribed — a right hotly contested during the public debate on the draft higher education act last year. A parallel situation exists in the medical colleges, where the Ministry of Health can intervene on admission numbers but where, in recent years, the ministry has been pressing for more admissions even though teaching facilities are inadequate.

Last September, the Warsaw Medical Academy took advantage of the then liberalization to reduce its student intake — and triggered a sit-in by the parents of students who had failed to gain admission. Significantly, however, neither the Ministry of Health nor Solidarity and its student adjunct, the Independent Students' Association NZS, backed the parents, on the grounds that the decision must rest with the academy.

Unemployment among young graduates has been a problem in Poland for several years, until the imposition of martial law on 13 December 1981, with its system of compulsory registration and direction to — if necessary — unskilled manual labour.

Part of the explanation is that student admissions have been uncontrolled. University tuition in Poland is free, textbook prices are low, and during the 1970s artificially pegged food prices made it relatively easy for all but the poorest families to support their children through five or six years of higher education. Recent moves to adjust the price structure have, however, produced increases so high that many citizens cannot afford to purchase their (at best scanty) food rations, so that keeping a son or daughter at university is fast becoming a luxury. A system of maintenance grants and/or loans is being considered, which would, presumably, allow the ministry to regulate admissions without having to decree a *numerus clausus* on admissions. It could also, in the long term, help to end the controversial system of "bonus points" by which young people from working class or peasant families receive extra marks in the university entrance examination to compensate for the lack of intellectual background at home. It could also (although this has not been mentioned publicly) provide the ministry with a further means of controlling student unrest without encroaching on university autonomy by ordering that offending students should be expelled or suspended.

Vera Rich

## British nuclear power

### Risks assessed

The British Central Electricity Generating Board (CEGB) published a final appendix to its case for building a pressurized water reactor (PWR) in the nick of time last week, just one working day before the public inquiry into the siting of the reactor at Sizewell in Suffolk opened this Monday (26 July). The publication of appendix M, on degraded core accidents and their consequences, added the last few kilogrammes to the more than 100 kilogrammes of reading matter that is CEGB's statement.

Participants at the inquiry, however,

### The Sizewell inquiry

Organizing the Sizewell inquiry is by no means simple, to judge from the length of meetings set aside for the task. The hearings earlier this week were the second in a series of three to discuss the precise procedure of the main inquiry, beginning on 11 January 1983.

The Department of Energy has promised a wide-ranging inquiry that will consider the need for the proposed pressurized water reactor (PWR) in the light of the government's long-term energy policy, the safety of the design, waste management and local environmental issues. Sir Frank Layfield, the inspector, favours hearing evidence topic by topic rather than each organization making its complete case in one session as at the Windscale inquiry. The order of witnesses is expected to be the Central Electricity Generating Board (CEGB) followed by government departments and other statutory bodies, local authorities, objector groups and individual objectors, although the order may well depend on the topic.

The procedural hearings are designed to help the inspector to determine precisely how to order the main inquiry — for example, how to split topics and what guidelines to issue on the release of documents. The formal opening of the inquiry this week gives the inspector statutory powers to request documents he believes should be made available before the main hearing. The inspector seemed eager to take up those powers on Monday when he promised to draw up guidelines on the release of documents that have not been voluntarily disclosed.

The inquiry is planned to take place at the Maltings, Snape, the closest large meeting place to the proposed site. But the inspector has suggested that part of the inquiry, perhaps during June and July 1983, should transfer to London in view of the national importance of much of the debate. Some objectors will be arguing the case for holding more of the hearings in the capital, especially those on need and economics.

have longer than a weekend to read and digest appendix M. This week's opening of the inquiry was purely formal, designed to give Sir Frank Layfield, the inspector, statutory powers to request documents before hearing the first evidence on 11 January 1983 and to deal with procedural matters (see box).

Although making up the tail of CEGB's case, appendix M is by no means insignificant. It is concerned with the probabilities and consequences of major accidents resulting in the melting of the reactor core. CEGB commissioned studies from the Westinghouse Electric Corporation, the company from which the basic PWR design is being licensed, on the probability of degraded core accidents, and from the British National Radiological Protection Board (NRPB) on their radiological consequences.

The risk analysis is similar to that pioneered by Rasmussen in his 1977 study of fault sequences in PWRs. Thus the Westinghouse analyses are based on event trees which trace the probability of minor faults leading through a chain of further incidents to core meltdown. The NRPB study assesses the risk to individuals living near a reactor of each type to 12 types of release considered by the Westinghouse Corporation.

The errors in the analyses are themselves uncertain, says CEGB. Nevertheless, it draws conclusions. Degraded core accidents are expected to occur 1.16 times every  $10^6$  years, near enough to CEGB's own limit of frequency of  $10^{-6}$  severe accidents per reactor year. But 97.5 per cent of core meltdowns would not breach the containment, says the appendix, so that the estimated frequency of core meltdown accompanied by containment failure or bypass is  $3 \times 10^{-8}$  per reactor year. But CEGB admits that one of the most uncertain aspects of the studies is the frequency of core meltdown resulting in containment failure.

According to NRPB's analyses of the consequences of core degradation the risk of early death to an individual living near the reactor is  $10^{-9}$  per year, says the appendix. An accident resulting in no more than one death would be expected about every  $10^6$  years. But an accident involving 1,000–6,000 deaths, the largest number considered, would be expected only about once every  $10^{10}$  years. Similarly, the risk of cancer directly attributable to the reactor is found to be very low, 40 million times less than the general risk of contracting cancer.

CEGB believes these analyses are more refined and accurate than those used by Rasmussen. The probability of a core meltdown is within the bounds of acceptability, it says, although it is asking for further assessments of the uncertainties. Whether CEGB's confidence will convince those who are sceptical of the value of risk analysis must remain until 11 January to be seen.

**Judy Redfearn**

## Academy rebuts Idso on CO<sub>2</sub> research

Washington

A new National Academy of Sciences (NAS) study\* has upheld the conclusions of an earlier one in concluding that a doubling of carbon dioxide in the atmosphere will cause a warming of the globe's mean temperature by  $3^\circ \pm 1.5^\circ$  centigrade. The conclusion, which is based on computer models, came as no surprise, as a consensus on this point has been building among scientists for some time.

But the NAS study did raise eyebrows by singling out for rebuttal the views of Dr Sherwood Idso. Idso has been a vocal minority, of virtually one, in dissenting from the generally accepted view of the effects of carbon dioxide increases.

Dr Idso, a physicist at the Department of Agriculture's Water Conservation Laboratory in Phoenix, Arizona, asserts that his "natural experiments" show that the prediction of the modellers is too high by an order of magnitude. The NAS study attempts to rebut his conclusions by showing how Idso's experiments are "on time and space scales clearly inappropriate to the carbon dioxide problem and do not involve the components of the climate system that are important for long-term climatic change".

The panel was apparently split over whether to single out for criticism Dr Idso (and, to a lesser degree, two other

researchers, Drs R. E. Newell and T. G. Dopplick). A minority on the panel had felt, as one participant put it, "that we shouldn't dignify the arguments of Idso with a comment". But the panel eventually decided that Dr Idso's findings needed to be answered to set the record straight. "Policy makers were getting confused", says Dr Joseph Smagorinsky of the Geophysical Fluid Dynamics Laboratory at Princeton University, the chairman of the academy panel. Dr Stephen Schneider of the National Center for Atmospheric Research, whom the panel had called in as an "invited expert", pushed hard for an official rebuttal of Dr Idso's conclusions. "I think there was a period of about six months when my phone rang every hour with someone wanting to know why" there was this difference of opinion between Dr Idso and everyone else, he says.

As for Dr Idso? "I am a very small minority sill," he admits. But he appears undaunted: "The model results are far from conclusive". He says the modellers themselves admit their inability to include many important factors, for example cloud feedback and aerosol effects. "When you read their caveats, it seems to me that anyone who put faith in the models is foolish, really."

**Stephen Budiansky**

\*"Carbon Dioxide and Climate: A Second Assessment", National Academy Press, Washington, DC.

## French robotics and engineering

### More, please

Paris

Jean-Pierre Chevenement, the French minister of state for research and industry, fired a broadside at his critics last week in a major speech outlining his policy for the robotics and mechanical engineering industries. "I am not a technology fanatic", he said. "We do not choose technology; it imposes itself on us." France must compete in the world economy and so must automate and modernize as fast as her competitors. "France is not on the Moon; it is on Earth."

Chevenement said that change would have to be fast. The productivity of French industry would have to increase by 7 per cent a year for the next ten years, to double over the decade. The social consequences of this change would be vast: as great as the historic shift of labour from the land to the cities. But the change would be democratic: efforts were already under way to establish "contracts for progress", outline development plans for industrial sectors to be signed by industry chiefs, trades unions and government, which would take into account the fear of automation increasing unemployment.

In reality, the minister said, automation need not increase unemployment at all. Japan, through the automation of the car industry, had opened up new markets in

Europe and America, and so had increased (Japanese) employment in that area. France could do the same in other sectors, and could also increase employment in the industries that must supply the hardware and software (for example robots and control systems) which will re-equip the rest of French industry.

Chevenement has even introduced a new word to the French language to describe these industries: "*productique*", a word that — for the moment — lacks the resonance of unemployment and de-skilling which attaches to "robotics" and "automation". *Productique*, as Chevenement defines it, is the industry of advanced (typically electronically controlled) machinery, robots, industrial software, computer-assisted design and systems engineering. It employs 20,000 people and has a total turnover of FF8,000 million (£700 million) (see Table). This group of industries, together with electronics, will transform the whole of the rest of French manufacturing industry. It will "reduce human intervention" in systems of production. It will also "modify the balance between capital and labour, and increase the role of intelligence as a productive factor". Thus "tens of thousands of workers" must retrain, and means must be found to re-educate them and to shift France's system of professional education towards new forms of employment.

But for the present, the state of auto-



mation and the productivity of French industry were "disturbing". Management had often preferred to avoid the risks of industrial conflict over automation, and capital investment since 1974 had been small. The result was an "obvious under-equipment" of French industry. For example, France has only half as many robots per worker as Sweden or Japan, and the mean age of industrial capital equipment had increased in France from 14 years in 1974 to 16 years now.

What's to be done? As usual, M. Chevenement prescribed massive financial intervention by the government: FF2,500

million (£200 million) over 3 years for the machine tool industry (following a plan established by his predecessor at the ministry of industry), the launching of a robotics research programme involving 300 researchers and other such grand concepts.

Last week, however, a new note was struck. Part of the investment must flow from a reflation of the home market, said Chevenement, who thus offered a strong hint to the finance minister Jaques Delors that the recent freeze on French wages must not last too long. "The issues at stake are so important we cannot waste time" he said.

**Robert Walgate**

## Win for whales

Conservationists are jubilant. Last week the annual meeting of the International Whaling Commission (IWC) voted 25-7 in favour of phasing out commercial whaling by 1986. In view of the uncertainty over the numbers of whales and condition of stocks, the decision was taken to be safe rather than sorry. Meanwhile whaling nations now have a three-year breathing space in which to decide whether to accept the ban or to go their own way. Interim quotas have been allocated and fishing will continue until 1986.

Obstacles to an effective ban still remain. The motion, moved by the Seychelles, did not call for a ban or moratorium by 1986, but for zero catch limits. If new scientific evidence were to emerge that stocks were in a better condition the decision could be reversed. Whaling nations have 90 days to lodge a formal objection which under IWC rules would allow them to carry on fishing whales. By taking up such a position, whaling countries would keep their options open and it is a likely course of action for the majority.

The vote on fishing quotas was taken at an extended session of the meeting. Peru is to be allowed to take 165 bryde's whales in the 1982-83 season from a population that best estimates put at 1,000. Spain, which voted for the ban, was given a quota of 270 fin whales for the phasing-out period, with a maximum of 120 per year. Some estimates put the stock as low as 800. Japan can take 450 sperm whales this season and 400 next.

Many participants felt that in view of the phasing out of whaling such quotas would not have a serious effect on whale stocks. Some conservationists, however, viewed the quotas as a sell-out. They had hoped that the conservationist countries would stick together not only to impose a total ban but also to push through protection for the bryde's, fin and sperm whales.

Although after years of campaigning the conservationists have persuaded the IWC plenary session to vote for what is in effect a total ban, the outcome depends entirely on the Japanese. The other whaling nations will be swayed by the stand Japan takes. If Japan carries on whaling, IWC will have no further control over catch limits. The whaling nations will simply set their own. While environment groups are growing in Japan the fishing industry remains a powerful lobby. It is by no means clear that the United States has the political will to impose unilateral fishing and trade sanctions against Japan in the event of a decision to continue hunting. There is no other restraint — except Japanese good sense.

**Jane Wynn**

Indices of world robot production, according to figures contained in a French government report on robotics released last week

| Country      | Annual production (no. of robots) | Cumulated production | Cumulated production of robots each worth > £12,500 | Current turnover | Cumulated turnover | No. of workers building robots |
|--------------|-----------------------------------|----------------------|-----------------------------------------------------|------------------|--------------------|--------------------------------|
| Japan        | 11,000                            | 43,000               | 4,750                                               | £65m             | £217m              | 3,750                          |
| USA          | 8,130                             | 19,000               | 3,800                                               | £82m             | £192m              | 3,420                          |
| West Germany | 1,600                             | 4,800                | 1,200                                               | £16m             | £58m               | —                              |
| Italy        | 1,300                             | 3,900                | 1,000                                               | £16m             | £49m               | —                              |
| France       | 1,037                             | 3,815                | 687                                                 | £16m             | £52m               | 838                            |
| Switzerland  | 800                               | 2,400                | —                                                   | £1.2m            | £4m                | —                              |
| Scandinavia  | 560                               | 2,060                | 1,600                                               | £16m             | £57m               | 700                            |
| UK           | 80                                | 300                  | 30                                                  | £0.5m            | £2m                | —                              |

## UK plant biotechnology

### ARC joins in

The British Agricultural Research Council seems well on the way to joining up with a new biotechnology company specializing in plant genetics. The new company is being organized by the British Technology Group, the product of the *de facto* merger a year ago of the National Research Development Corporation and the National Enterprise Board. The group is expected to provide about a third of the initial capital of the new company, in which a total investment of between £12 and £15 million is being sought.

The new company is thus closely analogous to the British company Celltech, established in 1980 by a group of city institutions in partnership with the then National Enterprise Board. Part of the intellectual capital of Celltech is an agreement with the Medical Research Council under the terms of which the company has the first refusal to exploit discoveries arising in council establishments.

While the agricultural proposal has been in the air since the beginning of the year, it seems to have come to life only in the past few weeks, with an expression of firm interest from the British-based international oil company Ultramar. The intention now is that the British Technology Group will be drawing up a firm prospectus and business plan for the new company, which should be formed before the year is out.

There appears to be little danger that the new company will conflict with Celltech, which has apparently taken a policy decision not to engage in plant

genetics. Celltech's chief executive, Mr Gerald Fairtlough, said last week that his company welcomed the proposed company linked with the Agricultural Research Council and thought there might be opportunities for collaboration.

At this stage, none of the backers of the proposed company is willing to speculate about the directions in which its research and development may lead, although it does seem to be understood that it will not venture into the veterinary field, in which Celltech has declared an interest. The British Technology Group, destined to be a shareholder in both companies, says however that competition between the two would not be unduly worrying.

As yet, the Agricultural Research Council has not seen a formal version of an agreement for its participation in the company, which is nevertheless likely to be for a limited span of time in the first instance. The Medical Research Council's commitment to Celltech was for an initial period of five years.

The new company will be principally concerned with the genetic manipulation of plants, in which the council's Plant Breeding Institute at Cambridge and its John Innes and Rothamsted stations are involved together with research groups such as that concerned with nitrogen fixation at the University of Sussex. But Dr Ralph Riley, the council's secretary, says that the new company will also use other techniques for producing new strains of plants, including the propagation of plants from single cells by "conventional" cloning techniques. He says that the new company will aim not merely to carry out research and development but that it will also market new products.



## European geophysics

# Traverse planned

Plans to involve geoscientists from all over Western Europe in a joint study of Europe's continental crust moved a step forward this month when a working group of the European Science Research Council all but finalized a proposal for the European Geotraverse project. The cost will exceed FF 22 million (£1.8 million).

The proposal will be considered by the European Science Foundation at its annual assembly in November. If it is approved, the 5–7 year project will be a test of whether the European science community can make such large-scale cooperation work.

The geotraverse will survey a swathe of continental crust extending 4,000 km from the northern tip of Scandinavia southwards to North Africa and will cover crustal ages ranging from the oldest, in the Precambrian areas of Scandinavia, decreasing southwards to the tectonically still-active Mediterranean region.

The traverse breaks down into three regions. The northern segment consists of the Baltic "Shield" and the regions just south of it, extending from the North Cape to north Germany. The Precambrian shield is between 600 and 3,600 million years old, while the neighbouring Caledonides in Scandinavia and north Germany are between 600 and 400 million years old. The Precambrian crust has remained essentially unaltered since the close of the Precambrian 600 million years ago and should contain evidence of palaeotectonic processes that may have taken place during the earliest stages of crustal development. For example, geochemical and morphological studies may yield evidence of early subduction processes beneath the surface, while palaeomagnetic measurements should reveal what, if any, plate motions occurred during this era.

South of the Baltic Shield lies a region of Precambrian basement overlain by a younger formation, several kilometres in thickness. It is hoped that deep seismic observations will show how the Baltic Shield extends into the basement and meets the younger Caledonide basement of north Germany.

The middle section of the traverse, from 400 to 230 million years old, coincides with the mountainous "Hercynian" area of Central Europe. Seismic mapping has revealed the complex structure of the upper mantle in this region, while several models exist for the tectonic processes leading to the mountain development. Deep seismological studies should demonstrate which models are correct, and petrological studies of young volcanic rocks should yield evidence on the composition of the deep crust. The substructure has been investigated by gravity measurements and the project may explain some of the anomalies discovered.

The southern section of the traverse

includes the Alps, the Po basin, the western Mediterranean and the north continental margin of Africa. This region has been dominated by Alpine mountain-building and, further south, by the relative movement of African and European plates. The combination of geophysical and geochemical techniques will, it is hoped, answer such fundamental questions as how the height of the Alps varied with time and what processes produced the western Mediterranean basins.

The geotraverse project will be coordinated by the European Science Foundation, but the expenditure on science will be provided by the respective funding bodies of the participating countries. Moreover, as well as the large-scale coordinated interdisciplinary experiments along the traverse, the project requires smaller-scale regional programmes to provide supplementary information. Thus, although the European Science Foundation may approve the project in November, it will be up to the individual countries to make the enterprise work. At present the signs are very favourable.

**Philip Campbell**

## Romanian agriculture

# Meat over bread

The Romanian leadership has this month published a grand design to reform the nation's eating habits and to eliminate nutritional diseases. But can Romanian agriculture meet the challenge?

The programme is nutritionally conventional. The targets for 1985 include a drop in calorie intake from the present average daily intake of 3,300 calories per head of population to 2,800–3,000, an emphasis on lean meat and dairy products and a reduction of the proportion of cereals in the diet from the present 45 per cent to about 38 per cent. Fat and sugar consumption is to go down, and the balance is to be provided by vegetables and fruit. By 1985, the plan is that the average Romanian will consume 80 kg of meat and fish and 230 litres of milk a year.

The programme gives no hint, however, of where the new diet is to come from. Food queues and shortages are endemic in Romania and agriculture is in disarray. Last year, President Nicolae Ceausescu castigated agricultural students who do not intend to work on farms but who simply see their diplomas as entrance tickets to the civil service. During the past three months, high-level reorganization of the agricultural ministries has been accompanied by veiled allusions to slacking and corruption.

The new programme reiterates the need for local self-sufficiency. Fortunately Romania's only major conurbation — Bucharest, with a population of 1.8 million — is prudently coupled for administrative purposes with the Ilfov farming district. In the event, food rationing was introduced last autumn, concurrently with self-sufficiency. The present rations (1 kg per person per month of meat, flour and sugar, and 1.5 litres of cooking oil) compare oddly with the high-protein diet prescribed for 1985.

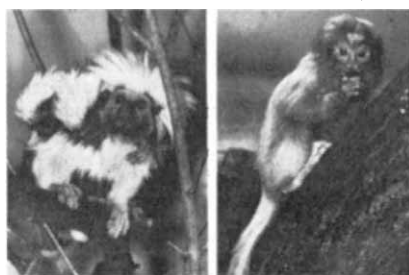
If the nutrition targets are to be met, some means must be found to reverse the falling trends of meat production. A scheme worked out by the Academy of Agricultural and Forestry Sciences may help. It advocates a complex system of research and planning in which multidisciplinary teams would work towards an integrated zootechnology to raise the quality and quantity of meat production.

The immediate bottleneck in Romanian food production, however, is in fodder and cereal production. In this respect, the reduction of human cereal consumption now proposed makes sense. Cereal production is not intensive — the target yield of 24 million tonnes for 1982 works out at less than 4 tonnes per hectare. A drop in bread consumption would leave more cereals available for animal feed. But the immediate outlook is one of tightened belts and severe rationing.

**Vera Rich**

# Appealing primates

The smallest monkeys in the world are alive and thriving in the Charles Clore Pavilion for Small Mammals at London Zoo. The new exhibition, "The World's Smallest Monkeys", which runs until the end of October (free entry for visitors to the zoo), is part of a campaign to attract visitors and to focus attention on small and distant primates, the marmoset and tamarin monkeys.



*A cotton-headed tamarin with twins (left) and a golden lion tamarin, all to be seen at London Zoo*

Their natural habitat, the tropical forests of South America, is rapidly being destroyed and 12 of the 15 different species are now seriously at risk. The zoo hopes the public will be made aware of these species of monkeys — the pygmy marmoset weighs only about 100 grammes and is difficult enough to spot in its zoo habitat — and also wishes to highlight its conservation work. The golden lion tamarin and cotton-headed tamarin are both virtually extinct in the wild but have bred successfully at the zoo. Both are potentially important laboratory primates. Their small size has both a practical advantage and, it is hoped, large public appeal.

**Jane Wynn**

# CORRESPONDENCE

## Social sciences

**SIR** — The leading article in *Nature* of 29 April (p.789) discusses the place of the social sciences *vis-à-vis* the natural sciences but does not pinpoint the essential differences.

Physical scientists base predictions on firmly established laws, and though individual scientists are far from infallible their answers have a good chance of being right. Engineers deal with situations for which there are no unique solutions — for example, a railway can be built along a number of alternative routes but once the choice is made the basic knowledge for its construction and operation is available.

Economics is different. Professors, politicians, bank economists, businessmen, unionists and treasury officials all advance remedies for economic ills. But no matter how confident these experts may sound, there are as yet no accepted criteria for evaluation. Even when a policy based on one line of thought has been selected the government is likely to abandon it half way through.

Still further from mathematical precision lie the social and behavioural sciences, including education. There is a great and pressing need for study and research, best based on a background of mathematics and statistics. But the variables are so numerous that seldom can rigorous laws be established.

Social scientists should have theories. Too often, however, these are advanced as facts (and no amount of repetition can convert an unproved theory into a fact!). Sociologists are at a disadvantage compared with applied scientists who, if basic information is lacking, can turn to controlled experiments. This is, I think, the essential difference. Social scientists must rely on inductive logic and judgment: the physical scientist can test his theories in the laboratory.

Were social scientists to acknowledge the limitations of their methods to the public and to advance their views as hypotheses rather than truths, would they not attract increased support?

IAN W. WARK

Melbourne, Australia

## A private matter

**SIR** — I wish to point out a blooper in your leading article, "Colleges back from the dead" (*Nature* 10 June, p.443). Seven thousand dollars would indeed be a stiff fee for a year's tuition at a state university in the United States. It is safe to say, however, that no undergraduate at a public school is, for the moment at least, faced with a tuition fee even approaching this amount. While, to the British ear, "University of Southern California" may sound like the name of a state university, and while USC's football team has repeatedly proved itself equal to, or better than, that of any large state school (much to the dismay of University of California fans, of both the Los Angeles and Berkeley persuasion), nevertheless USC is, contrary to your implied assertion, an undeniably *private* university.

DAVID L. GLANZMAN

Department of Psychology,  
University of California,  
Los Angeles, USA

## No "remote viewing"

**SIR** — A long drawn out controversy in *Nature* has followed the claim made by Puthoff and Targ<sup>1</sup> in 1974 that certain individuals can perceive objects or scenes blocked from ordinary perception by distance or shielding (the supposed phenomenon called "clairvoyance" by earlier generations). The most recent communications<sup>2,3</sup> raise issues about the content of Puthoff's and Targ's experimental records that can only be settled by direct examination of those records. However, their critic, Marks, has reported their refusal to grant him access to their records<sup>2</sup>.

In an attempt to clarify this issue, writing as a *bona fide* investigator of long standing in this general area but having no prior involvement with these particular experiments, I recently requested access to the data on the Price and Hamid series on which Puthoff and Targ based their original case for "remote viewing". No reply has been received after an interval of two months, despite repeated approaches. It must be concluded that the evidence offered by Puthoff and Targ is not accessible to other investigators. In this sense their claim can no longer be regarded as falling within the scientific domain, and further public discussion appears unnecessary.

CHRISTOPHER SCOTT

London N19, UK

1. Targ, R. & Puthoff, H. *Nature* 251, 602-607 (1974).
2. Marks, D. *Nature* 292, 177 (1981).
3. Puthoff, H. & Targ, R. *Nature* 292, 388 (1981).

## Centenary plea

**SIR** — At the meeting held in Cambridge last month to mark the centenary of Darwin's death, a statement was signed by 199 of the participants. The signatories ranged from Nobel laureates to graduate students and came from 20 different countries. They included specialists from the whole range of disciplines that impinge on the study of evolution. The statement read as follows:

We have gathered at a conference to commemorate Darwin and to discuss the evolution of plants and animals in the past. However, we are also deeply concerned about what may happen in the future. Human beings are creating conditions that could easily bring the long evolutionary process to an end with the total destruction of all life. In particular, the arms policies of the major powers are desperately short-sighted and increasingly unstable. Even if the threat to use nuclear weapons has acted as a deterrent in the past, it is not likely to do so for the rest of time. The continued spread of nuclear weapons and the development of new chemical and biological weapons could easily result in a war that rendered this planet uninhabitable. We urge our governments to take a long view of what they are doing. Human beings have evolved with an intelligence that has created technologies of enormous power. This intelligence must now be harnessed to secure a long-term future for life on Earth.

PATRICK BATESON

Department of Zoology,  
University of Cambridge, UK

## Index for patents

**SIR** — Patent specifications are one of the several forms of the primary sources of scientific and technological information. Announcements of patent specifications are primarily through the official gazettes of the patent office of a country, or through secondary services like *World Patent Index* and other abstracting and indexing services brought out by Derwent Publications Ltd, London. The official gazettes follow their own system of arranging the abstracts, while the abstracts in *World Patent Index* and in other secondary services follow the *World Patent Index* scheme. These services generally provide accession number, patent number and patentee index, but no subject index.

Searching would be made more simple and less time consuming if patent information services provided a subject index. The user is more often aware of the subject area in which he requires information than of the name of the patentee or patent number.

K.C. GARG

Defence Scientific Information  
and Documentation Centre, Delhi, India

## Help for refusniks

**SIR** — We were glad to see a letter of Professor Legay, Secretary General of the World Federation of Scientific Workers (WFSW) in *Nature* of 11 February (p.452). No doubt, the federation can help to solve the problem of emigration of "refusnik" scientists from the Soviet Union, if it really wishes to, especially in view of the respect and influence it receives in this country. But Professor Legay, perhaps unwittingly, downgrades the dimensions and the acuteness of the problem.

We want to emphasize that the question is not about several separate individuals but about a large group of scientists who are subsequently refused permission to leave the country and whose access to normal scientific activity is, at the same time, being deliberately restricted or almost completely blocked in certain areas (such as seminars, conferences, teaching and, to a large extent, publications). As we understand it, this is precisely the situation in which WFSW can and must intervene and help according to the goals expressed in its charter.

It would not be an exaggeration to say that most of us feel as if sentenced to professional death. But there is also a great psychological pressure that every refusnik, not necessarily a scientist, undergoes and that results from general uncertainty about the future, especially the future of children, from the impossibility of getting any information on how long we have to wait (and some of us must wait 10 years or more for permission to leave), and from the absence of legal ways to contest refusals and to defend ourselves. It has to be understood therefore that we speak not only about emigration as such but about our very existence and the existence of many others who are in the same predicament, our existence as scientists and human beings, about the lives and future of our families.

The situation continues to deteriorate (contrary to the hope expressed by Professor Legay at the end of his letter). All scientific seminars organized by refusniks were closed



by force, Victor Brailovsky, the host of the most well known of them, was sent to exile, a number of refusnik scientists recently lost their jobs. All of this indicates that the situation is approaching a critical stage which may become irreversibly tragic. As the most sinister sign we consider several recent statements of some officials that the problem of Jewish emigration from the Soviet Union no longer exists and that there are no more Soviet Jews wanting to emigrate (see for example the *Baltimore Sun*, 11 April).

Therefore we find it necessary to emphasize once again that emigration to Israel remains the only possible solution to our problems and that our decision to emigrate remains unshakeable.

So far the activity of Professor Legay has been limited to asking for necessary information only from the official Soviet Trade Union organization connected with the federation. In this way, a very serious human problem is likely to be reduced to a formal argument the outcome of which very much depends on the knowledge and arbitrariness of some bureaucrat or other. We are convinced that the problem can be properly understood only through direct contacts with those who are personally involved. We use this opportunity to invite Professor Legay or another representative of WFSW who would be able to consider the problem honestly and comprehensively to come here and to meet us and other refusnik scientists.

If WFSW wants to help us, it can do so by giving the problem the widest publicity. At this time it may seem that the problem is ours only, and, in a narrow sense, it is. But the solution is far-reaching as it may affect (and already does affect) the lives of many scientists all over the world, their interaction and their confidence in each other.

SOLOMON ALBER  
JACOB ALPERT  
EITAN FINKELSTEIN  
GREGORY FREIMAN  
ALEXANDER IOFFE  
JOSEF IRLIN  
ARKADY LEBONOV  
ALEXANDER LERNER  
NAUM MEIMAN  
MARK REITMAN

Moscow, Soviet Union

SIR — We are very concerned with the impression that may have been conveyed in the letter to *Nature* (11 February, p.452) from J.M. Legay, Secretary General of the World Federation of Scientific Workers.

In discussing an earlier letter written by a group of refusniks, he seems to have the impression that the number of scientists who have requested and been refused emigration visas is comparable to that of the number of signatories. He conveys this impression by comparing this letter with one he had received personally and suggests that "the situation has evolved since last year". The implication is that some of the signatories have emigrated. If this is the message, it is erroneous as well as dangerously misleading.

In a recent visit to the Soviet Union, two of our members ascertained that none of the signatories, either of the private or the published letter, has been granted an exit visa. Furthermore, our emissaries were assured that the problem is orders of magnitude greater than Mr Legay suggests. In Moscow alone, there are at least 50 Doctors of Science who are refusniks in addition to 300 Candidates of

Science, among an estimated total of 1,000 scientific workers in the same predicament. For the whole of the Soviet Union, with refusniks in Kiev, Leningrad, Kharkov and so on, a reasonable estimate would be a minimum of 2,000. An indication of the situation is a letter from 46 scientist refusniks which we hope will soon be published. These scientists, like their ten colleagues, to whom Mr Legay was replying (*Nature* 24 December 1981, p.688), have described the tragedy of their plight better than we could possibly do.

In the light of the fact that many refusnik scientists and their families have been subjected to arrests, long-term imprisonment, revocation of degrees (*Nature* 6 May, p.4), and other forms of harassment, it is remarkable that they continue to speak out forcefully.

ARTHUR YELON  
ADI EISENBERG

Committee of Concerned Scientists,  
Quebec, Canada

## More on Ovid

SIR — I dare say your correspondent M. Kamen-Kaye (*Nature* 8 July, p.114) would find much of interest on the early history of geology in Lyell's *Principles of Geology* (London, 1830). Ovid is quoted, as are other classical authors, and whatever one may think of Lyell's abilities as a historian, he was a damn good writer.

GORDON CHANCELLOR

University of Oxford,  
Department of Geology and Mineralogy,  
Oxford, UK

## Safe harbour

SIR — Sr Alberto C. Taquini's letter regarding the Falkland/Malvinas Islands (*Nature* 10 June, p.430) omits the following salient facts.

- (1) Only through British help did the Argentine manage to throw off the Spanish yoke.
- (2) After 1829 the Argentines used the islands as a convict settlement, the convicts revolted and made the Governor a prisoner.
- (3) The intervention of the captain of HMS *Cleo* was due to this fact, and interference by the Governor and/or the convicts with an American whaling ship.
- (4) Port Stanley was a port of refuge for sailing ships damaged in the passage round Cape Horn, ships of all nations. In fact the only port of refuge available for many miles.
- (5) It was clearly to the benefit of all that a stable administration was established there.
- (6) One can only imagine the Governor was thankful to be rescued and returned safely to his homeland as his country was either unable or unwilling to restore law and order.

JOHN R. BARTON

Winsford, Cheshire, UK

## Falklands alone?

SIR — If F.W. Cousins would consult the mass of evidence assembled and sifted by Dr J. Goebel in *The Struggle for the Falkland Islands* (Yale University Press, 1927), he would be forced to agree that the account of the history given by Señor Taquini is broadly correct (*Nature* 10 June, p.450 and 1 July, p.8).

The early navigators' accounts are far too confused for any certainty about the first sighting, though the most likely candidate is a ship of Camargo's armada in 1540, (note the

mention of foxes — Darwin's 'Falkland fox', *Canis antarcticus*?). But in any case discovery by itself has never been accepted as conferring sovereignty (did not Cook discover New Caledonia?).

The British case rests on the technicality that the establishment of Fort Egmont on West Falkland in 1766 predated the transfer of de Bougainville's colony on East Falkland to the Spanish crown in 1767. The French could have held out for their prior rights, but in view of the close relationship between the Bourbon kingdoms of France and Spain they backed down. It seems that the abandonment of Port Egmont was in fulfilment of a promise by George III, and it would hardly be honest to pretend that because it was unwritten it was never given.

Of the events leading up to the British takeover of 1833 it must be said that Louis Vernet's colony was broken up in 1831 by an American corvette, for alleged piracy, that the last Argentine governor was murdered by the remaining settlers shortly after his arrival in 1832 and that the British in 1833 found a state of anarchy on the islands. Vernet accepted compensation of £2,400 for his lost colony. Still it must be admitted that in 1833 the British claim to West Falkland was weak and to East Falkland weaker.

The important question, though, is why we are still talking about the rights and wrongs of events that happened before living memory, that affected very few people and that decided the fate of a very small territory. The restraint of Argentinian governments before Peron was more fitting to the dignity of a great nation than the subsequent strident claims.

Whatever happened up till 1833, circumstances now are wholly different. Half a dozen generations of British people have lived in the islands, which are now no more Argentinian than Normandy is English or Mexico Aztec. There can be no going back to 1832. However, it is not possible either to go back to 1 April 1982, now that a thousand people have died for the islands. Neither side is likely ever to accept the other's sovereignty. By elimination, the only solutions likely ever to be satisfactory are independence or permanent UN trusteeship. In view of the fact that it was the United Nations that introduced the notion of negotiable sovereignty in 1965, only independence seems likely to be acceptable to the islanders. Needless to say, independence would have to be declared by the islanders, rather than granted by the British, in order to be recognized by Argentina.

The question arises, how could so small a community with so few resources afford independence? The population would in fact only have to be multiplied by a factor of four to be equal to that of Nauru or Tuvalu, and this would not seem improbable. But biologists must view with anxiety the economic development that might go with population increase, for the islands are a laboratory and museum of sub-antarctic flora and fauna. Could not the international scientific community help to provide the dowry of an independent Falklands by agreeing to locate there the headquarters of organizations concerned with Antarctic research and conservation? The provision of services rather than the exploitation of natural resources would be the most appropriate basis for the islands' economy.

PHILIP J. STEWART

Department of Forestry,  
University of Oxford, UK

# Gene therapy

Bob Williamson\*

**Gene therapy is not yet possible, but may become feasible soon, particularly for well understood gene defects. Although treatment of a patient raises no ethical problems once it can be done well, changing the genes of an early embryo is more difficult, controversial and unlikely to be required clinically.**

In itself, genetic manipulation is nothing new; in fact, some anthropologists mark the dawn of civilization from the point at which man began to use selective breeding for crops and domestic animals. Eugenics, as applied to humans, was a powerful ideology during the first thirty years of this century, and appealed to a surprising diversity of geneticists, of all political persuasions. Only the rise of totalitarianism during the 1930s ended this enchantment, making gene therapy the highly charged and emotive phrase that it is today.

Why has human genetics again become the focus of public and scientific attention? The reason is that, largely as a result of recombinant DNA technology, genes have ceased to be theoretical concepts and have become DNA sequences, to be isolated, studied and changed in the laboratory. It is now *relatively* routine to isolate a single gene coding for a particular protein from the several million 'gene-size bits' of DNA in the nucleus of every human cell, and to examine its functions. More important, it is possible to isolate and compare the equivalent genes from different individuals, making it feasible to identify abnormalities caused by mutations of the DNA. Once a serious gene defect can be identified in an individual, two possibilities arise. If the identification is made during the first few weeks after conception, the birth of affected individuals can be prevented, or, in theory at least, an attempt can be made to replace the defective gene by a functional one. While the former is an acceptable and effective practice in many societies, there remains much interest in the prospects and problems of the therapeutic approach.

## Diseases for therapy?

The most obvious human genetic diseases which are candidates for potential gene therapy are those in which a single known gene does not function properly. In a few such diseases, altered DNA sequences have been cloned into bacterial plasmids. These can be used to identify whether an individual is homozygous for the disease or a carrier, and are disease-specific cloned probes. In this category are sickle cell anaemia and thalassaemia — in both of which the globin genes are defective — growth hormone deficiency, some of the

collagen disorders and Lesch–Nyhan syndrome in which there is an absence of the purine synthesis enzyme hypoxanthine phosphoribosyl transferase. In the near future, haemophilia (a defect of blood clotting factor VIII) and phenylketonuria will be added to this list, since attempts to isolate a gene probe are now being made in several laboratories. Further away, but equally interesting from the point of view of research (and the patients), is the possibility of determining the genetic defect of diseases such as Duchenne muscular dystrophy and Huntington's chorea; in these cases no defective gene product has yet been identified, but recombinant DNA technology allows one to bypass the product and hunt directly for an abnormal gene.

Single gene defects are important in clinical medicine. At present, in developed countries, they account for a high paediatric health-care burden — this is because they are chronic and often appear at birth, and may require replacement therapy with special diets or expensive natural products, such as blood clotting factor VIII or whole blood. In parallel with this clinical, economic and social perspective, there is also the personal problem faced by each family with a handicapped child. According to a recent World Health Organization report, there are over one hundred million carriers for potentially lethal inherited blood diseases such as sickle cell anaemia and thalassaemia, and approximately 200,000 affected children die each year from this single set of gene defects. In Britain, serious single gene defects (cystic fibrosis, muscular dystrophy and haemophilia are the most common) affect just under 1% of children.

## Termination or gene therapy?

It is now possible to carry out antenatal diagnosis for some single gene defects (and also for chromosomal defects such as Down's syndrome, as well as for neural tube defects), and to offer the parents a termination of pregnancy if the fetus is affected. It may shortly be possible to do this during the first 2 months of pregnancy, rather than after 18 weeks as at present, by using samples of fetal chorion. As the number of disease-specific DNA probes increases, antenatal diagnosis by DNA analysis or linkage should become available for all single gene defects<sup>1</sup>.

Why, then, even consider gene therapy for a life-threatening or handicapping inherited disease? Most communities at risk accept antenatal diagnosis and abortion of affected fetuses at the moment, as being better than having a severely handicapped child<sup>2</sup>. This is true even in countries where one might expect a resistance to abortion, provided that the community knows the disease and its consequences. But there are a few families (rather less than one might expect) who refuse termination, are misdiagnosed or who do not come for diagnosis in time because of a failure on their part or on the part of the system. Also, there is a range of single-gene inherited diseases that require a lot of clinical attention but are perhaps not so handicapping that the families (and society) regard termination of pregnancy as acceptable (haemophilia, or phenylketonuria, might be in this group in some countries). Although the decision to elect abortion rests with the parents, it does so in a social and legal context determined by all of us as to the degree of handicap that is compatible with a normal life for the child or the family.

## The route to gene therapy

If there is an identifiable need and demand for gene therapy (and the extent to which this is so will be discussed later), can it be done? At the moment, the answer is *unequivocally* no. This is an extremely important point; no gene replacement that is meaningful in the context of human genetic disease has yet been achieved even in model systems with animals.

Nevertheless, the way in which gene therapy might be carried out either on early embryos or on individual patients is clear in outline, from research now in progress, much of which was summarized at the meeting which inspired this article<sup>3</sup>.

## Patient therapy

Gene therapy of a patient must involve the insertion of a functional human gene into the relevant tissue, where it can respond to the normal physiological stimuli of the body. First, the defective gene must be identified and a normal counterpart isolated. A method must be found to put the normal gene into a cell where it can function (the primitive red blood cell in the case of sickle cell anaemia, for instance, or the liver for phenylketonuria). To obtain correct gene action, it may be necessary to put it into the correct site on the host cell chromosome, or even to delete the defective gene. It is our inability to obtain correct gene function when DNA is put into a cell, and the fact that few inherited diseases affect only a single tissue such as bone marrow, that makes gene therapy impracticable at this time.

It is only 5 years since the first cloned human genes, those for globin, were isolated. Today, it is a relatively routine matter to isolate any gene when there is a specific antibody, a portion of the amino

\*St Mary's Hospital Medical School, University of London, London W2 1PG, UK.



acid sequence of the protein encoded by the gene (for reverse synthesis of a bit of DNA using the coding rules, which can be used to pair with the natural gene), a tissue that selectively expresses the gene or a biological or drug selection assay for the gene. Each of these techniques has been used during the past year to identify genes that are expressed only at low levels in human tissues.

An isolated, cloned gene can either be directly injected into a cell or be carried in by a virus to which it has been linked by recombinant DNA techniques. Once in the cell, it may become integrated into the nuclear DNA, or remain free in the cytoplasm as a self-replicating, extrachromosomal element. In either case, it is usually transcribed to give RNA at the level of only a few copies per cell, as compared with the several thousand often achieved for endogenous genes. Translation may also be inefficient. Inserted genes are more stable if they are attached to another gene which can be selected, perhaps by drug resistance, and much effort is going into making virus-type gene vehicles with such selective markers which also cause the gene to be inserted into specific active sites on the chromosome.

### Embryo therapy

Other experiments have attempted to put genes into early embryos to study their integration and inheritance. So far, the experimental aims have been academic rather than practical, but there is no reason in principle that this approach to gene therapy would not work, in conjunction with *in vitro* fertilization. Indeed, it might well have a greater chance of practical success than treating a particular tissue, but it also raises greater ethical problems because it would affect not only the individual but also his or her offspring.

Genes can be inserted into embryos with viruses, or directly by microinjection. In either case, a fertilized egg (from rabbit or mouse) takes up the new genes and integrates them in a random number of linked copies into its own DNA. The site of integration appears to be random, and in most cases so far, the integrated genes have been inactive. But in a few animals, genes inserted at the embryonic stage seem to have been switched on in a tissue-specific way. The best evidence so far concerns the gene for metallothionein, which binds to heavy metals. Even in this case, and certainly for globin, it is not clear that the tissue where the gene is expressed is always the right one.

Such new genes have been inherited in a normal way through at least three generations. (Note that this is so even though the inserted genes come from a species other than that of the embryo; the animal cannot tell the difference once integration into the chromosome has occurred, but the experimenter can — whence the mixing of species). These experiments have so far all been carried out

on mouse and rabbit embryos, and early human embryos available from work on *in vitro* fertilization are not known to have been used.

In summary, it is possible to isolate normal genes, to put them into mammalian cells at any stage, including the fertilized embryo, and to obtain gene function in a random way and at a low level. However, it is not yet possible to ensure that the 'therapeutic' newly inserted genes function under normal control in the animal, in time, space or quantity.

### Is there a need?

To carry out gene therapy on an early embryo, one must be sure that it is affected by the disease — there is only a 25% chance that this will be true for most embryos at risk, and it might be thought that the parents, once the diagnosis has been made, would prefer to start conceiving again than consider gene manipulation of the embryo. Frankly, the prospect of gene therapy at the embryonic stage strikes me as ridiculous as a form of clinical therapy for a couple who have a 75% chance of a normal child (100% if they accept antenatal diagnosis) using the most popular, acceptable and enjoyable ways of having children that have been in vogue for many years without help or advice from molecular biologists. The correction of a disease by gene therapy will be worthwhile only if there is no other, simpler and more effective technique available.

On the other hand, treatment of tissue from a patient with a genetic disease raises few ethical problems, but is even more complex. The tissue must be accessible for isolation or transplantation, the gene must target to a region of the chromosome under appropriate control, and this is best done during early infancy, when the immune rejection system is still not fully developed. This must be balanced against the other options of antenatal diagnosis and conventional postnatal therapies, both areas where improvements are occurring rapidly.

### The thalassaemia model

Consider thalassaemia, one of the most common single gene defects, about which we know a great deal. This can be diagnosed antenatally by fetal blood sampling and sometimes by gene analysis. Carriers are usually detected in school, before they start having children. In communities at high risk such as Sardinia, Cyprus and the relevant immigrant communities in Britain, the incidence has fallen to next to nothing by a combination of screening and antenatal diagnosis. The communities involved want and demand healthy children; at the moment this means accepting, reluctantly, antenatal diagnosis and termination of affected pregnancies.

Even in countries best equipped for diagnosis, a few cases are born each year. In the Cypriot community in Britain, there were 2 known thalassaemics born last year,

compared with about 30 who would have been born without antenatal diagnosis. There is a special problem in the United States, where the campaign against abortion has led to campaigns against antenatal diagnosis even for crippling disease. So even in the most advanced countries, there are still some cases of thalassaemia. Are they candidates for gene therapy?

At first sight the prospect looks attractive. The disease is well understood, the normal gene is available and widely studied and its position in the chromosome is known. Perhaps most important, the tissue affected by the disease, bone marrow and blood, is available and can be treated *in vitro* and transplanted back into a patient. This treatment could be carried out after birth, and without the additional problems of dealing with embryos and altering their genetics. However, even in this best case, problems will arise, for the proportion of early precursor red blood cells in the marrow that could be taken from the patient, treated and reintroduced, is only a tiny proportion of the total stem cell population, and marrow transplantation still involves major risk.

### Other inherited diseases

In genetic diseases other than those affecting the blood, many tissues are affected. One obvious example is cystic fibrosis, the most common single gene defect in Britain, where lungs, pancreas, gut and sex organs can all be damaged to a variable extent. Even with dramatic advances in transplant therapy, it is unlikely that the removal of several tissues for individual gene insertion will become a realistic approach. In the absence of a 'target tissue', 'embryo therapy' has to be the chosen route, and for the foreseeable future this will remain unrealistic even if possible.

The entire question of gene therapy must also be seen in the context of rapidly changing biological and clinical knowledge. Personally, I think that there may be occasional cases where gene therapy is indicated (always in the future, when it can be done), and that in these cases we may well alter the inheritance ('genotype') of the individual as well as the way his body works ('phenotype'). This might be true for a period of a few years but I am certain that eventually far more sophisticated methods of treatment, probably involving therapy rather than gene alteration, will be developed.

### Polygenetic disease

In developed countries, the major health hazards (apart from those due to growing old) — coronary heart disease, hypertension, early psychotic illness, diabetes, some forms of cancer — have both genetic and environmental components, perhaps in roughly equal amounts. Indeed, there is doubtless a genetic component to our general physical

and mental make-up, although no one knows yet how important it is compared with environment. There is popular concern that gene therapy will be applied in this area, to alter the vigour, personality or 'intelligence' — whatever that may be — of an individual. The Council of Europe recently reflected this in a recommendation which requested provision "for explicit recognition in the European Human Rights Convention of the right to a genetic inheritance which has not been artificially interfered with, except in accordance with certain principles which are recognised as being fully compatible with respect for human rights . . ."<sup>14</sup>.

Such fears clearly strike a deep chord of response. Certainly, the hereditary factors in complex diseases such as diabetes and hypertension will be understood very much better as a result of research on isolated human genes. It is now possible to analyse, chromosome by chromosome, gene by gene, the components that make up any measurable inherited trait<sup>5</sup>. It has been postulated that a remarkably small number of cloned genes could 'cover' the entire human DNA with markers, and this seems to be true in practice<sup>6</sup>. The human gene map should be completely defined during the next 10 years.

Consider any serious illness, such as coronary heart disease, common in the community, a major killer of men aged between 30 and 50 in developed countries, with a large hereditary factor in this age group. There are very few cases where a single gene determines the risk; it is usually spread between several, especially those that determine the level of circulating cholesterol in the blood. Eventually, however, it should be possible to determine those at genetic risk of early heart attack soon after birth using DNA analysis.

Would gene therapy be possible in such cases? Since we know that many genes contribute to a healthy heart, it is unlikely that changing any one of them would lead to a dramatic physiological effect on the individual. Introducing genes would be difficult, and difficult to justify in the absence of a certainty of serious disease. Knowledge of the genetic factors is more likely to lead to new pharmacological approaches, and to prevention, than to gene therapy.

When we come to such vague concepts as 'aggression' or 'intelligence' or 'good looks', we are off in the land of the fairies. If there are any inherited determinants for these attributes, we are so far from understanding, isolating or measuring them that it is barely worthwhile even considering that an attempt at gene therapy could be made with any prospect of success.

## Ethics

Given that it is unlikely that gene therapy will be attempted on any but the smallest scale, and even so not for some time, is it necessary to discuss the ethics of gene therapy at all? I think so. There must be

two clear starting points, both of which are fundamental to most humanistic and religious views. The first is that any person, whether well or ill, deserves respect as an individual. This concept underlies the prohibition on experimentation on patients for its own sake — experimental procedures can be carried out only when there is hope that the patient, or his immediate family, will benefit, unless the procedure involves minimum risk (when full consent is still required). Patients, however ill they may be, cannot be used as laboratory animals, because this is a denial of fundamental humanism. This is why there are special safeguards for those at high risk — the psychiatrically ill and mentally retarded, the terminally ill and long-term prisoners.

When, in a much-publicized incident in 1980, Dr Martin Cline of the University of California at Los Angeles attempted to treat marrow from two patients, one in Italy and one in Israel, with normal globin genes, and then carried out a limited marrow self-transplant, he violated this precept (see ref. 7). Therefore, Cline's experiments were *fundamentally* unethical. His own work on mice showed that there was no basis for hope that globin gene insertion into marrow cells could give clinical benefit at that time. The thalassaemic patients were irradiated at a high level to clear the marrow in the long bone of one leg, a procedure of some risk, and they and their families were given hope that the gene therapy might help them in their fight for survival. It is unacceptable that patients should be misled in this way, and certainly makes it more difficult for medical research workers to argue that our house is in order ethically.

Cline can also be criticized for *procedural* reasons — he carried out experiments on patients abroad which he could not perform in his own hospital in Los Angeles, in countries with inadequate ethical committee control, and he did not notify authorities that recombinant DNA was used. However, these breaches, although real and of concern to me — the clinicians involved in Italy and Israel must examine their own consciences and decide whether they behaved correctly and with full knowledge of the proposed treatment — are far less important than the fundamental breach of ethics discussed above. It should also be noted, lest we in Britain become too smug, that hospital ethical committees are not legally required in Britain, and in the many hospitals where they do not exist, there is little effective control over experiments on patients apart from a reliance on the good sense and judgement of the doctor involved.

The second underlying ethical 'rule' which most (but not all) people would accept is that the medical profession, as a caring profession, should do all it can to prevent serious handicap. Scientists who work with clinicians share this rather difficult obligation, and society has

decided that part of it is that a termination of pregnancy before approximately 5 months is allowable if the child would suffer a serious handicap. These are shifting sands, and the definition of 'serious handicap' is a social one, and rests with the community and the family at risk, and only secondarily with the doctor (and even less with the scientist).

It is often asked whether gene therapy poses long-term genetic problems to human inheritance. Naturally, this could arise only were the early embryo treated, so that the newly introduced genes are inherited; treatment of a patient's marrow or liver could not lead to genetic alteration. The human genome contains many tens of thousands of genes coding for polypeptides, and many hundreds of thousands of pieces of DNA of unknown function. Each of us 'normals' carries 20 or so potentially lethal recessive genes, from whose lethality we are protected by also carrying a healthy dominant version of the same gene. Our children will inherit some recessive lethal genes from each parent, but since our mates almost always carry a different set of deleterious genes, inherited disease is unlikely. Gene therapy, at worst, could inadvertently introduce an extra recessive lethal gene to an individual, increasing the load from about 20 to 21, in return for which a person who would otherwise be chronically ill, becomes healthy and a normal member of the community. From the point of view of the clinician there seems to be no fundamental ethical problem here.

## Conclusions

The ultimate decisions as to whether and how we use gene therapy will rest with the community. Research with model systems and cell culture will continue in any case; these are well established and have other obvious benefits both to science and medicine. They will explain how normal and abnormal early embryonic development occurs, what causes miscarriage and chromosome defects, and the effects of environmental agents such as rubella or thalidomide.

Gene therapy is clearly not on the agenda for next month or next year. In my view, experiments with patients are clearly unethical at this time. When gene therapy is available, there must be a clear distinction between treatment of a serious disease in a way that affects only the patient himself, and treatment of an early embryo so that its inheritance is altered. However, gene therapy of both types will be possible in the future, and it should be considered now, before the headlines break on us all.

1. Orkin, S. H. *Nature* **296**, 202–203 (1982).
2. Williamson, B. *Nature* **292**, 405–406 (1981).
3. *Prospects for Gene Therapy — Facts and Fiction*, (ed. Friedman, T.) (Cold Spring Harbor Laboratory, New York, in the press).
4. *Council of Europe Parliamentary Assembly recommendation 934* (1982).
5. Ruddle, F. H. *Nature* **294**, 115–120 (1981).
6. Williamson, B. *Nature* **293**, 10 (1981).
7. *Nature* **291**, 369 (1981).



## NEWS AND VIEWS

# De-fogging cloud determination algorithms

from Ann Henderson-Sellers

THE International Satellite Cloud Climatology Project (ISCCP) is designed to take advantage of the global coverage provided by the current array of meteorological satellites. The project is planned to begin early in 1983 and algorithms to determine cloud cover, type, height and optical properties from satellite information are now being designed. At a recent meeting in Ottawa\* cloud analysis algorithms from the different groups participating in ISCCP were tested and compared.

Ultimately, ISCCP hopes to produce a comprehensive global cloud climatology. This should permit the validation of explicitly calculated cloud fields from general circulation climate models, improve the parameterization of cloud-radiation feedback schemes in simpler climate models and provide test case data with which to compare sensitivity studies involving critically important cloud-climate interactions. The project should also make it possible to assess the probability of obtaining cloud-free viewing conditions from space for any particular area.

Two important phases of the project will follow the comparison of cloud analysis algorithms. Interim cloud data and original radiance data must be gathered, reduced and archived so that alternative cloud conversion techniques can be applied as they

become available. Following this, cloud algorithms will need to be validated against traditional surface-based observations.

Determination of cloud characteristics from satellite information is extremely difficult and it is even possible that a unique solution will not be possible with data of the current resolution. The deter-

3.7  $\mu\text{m}$  waveband channel available from the Advanced Very High Resolution Radiometer on the NOAA polar orbiter satellites to identify cirrus cloud (A. Arking, NASA Goddard Space Flight Center). The effects of broken clouds, which are often smaller than the resolution of the sensor, can be modelled but it is not

yet possible to invert these results to determine uniquely the characteristics of the cloud field.

To facilitate comparison of cloud algorithms at the workshop the Chairman, Bill Rossow (NASA Goddard Institute for Space Studies), had designed answer sheets to tabulate and graph results in a uniform framework.

All the participants in ISCCP had been asked to analyse satellite data from three areas that represent difficult situations for cloud analysis procedures (see the figure). The chosen regions were (1) the northeastern US containing snow-cover, travelling cyclones, cold air outbreaks and the land-ocean interface; (2) the north-east coast of Brazil containing part of the inter-tropical conver-

gence zone, diurnal variation of convection over land and trade wind cumulus; and (3) the west coast of Chile with large diurnal variations of temperature over the land, many stratocumulus clouds and a few travelling storms.

*Ann Henderson-Sellers is in the Department of Geography, Roxby Building, University of Liverpool, PO Box 147, Liverpool L69 3BX.*

1700 08FE73 11A-4 00101 19111 WC1

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

Satellite picture taken at 1700h on 8 February 1979 showing cloud cover. (Reproduced with permission of NOAA/NESS.)

mination of cirrus cloud and of small broken cloud is particularly important. Climate models suggest that cirrus cloud affects climate in an opposite manner to other clouds — increasing the amount of cirrus cloud raises both surface and atmospheric temperatures (that is, increases the 'greenhouse effect') whilst increasing lower level cloud amount lowers temperatures. It may be possible to use the

\*The ISCCP Cloud Algorithm Intercomparison Workshop was held at the COSPAR meeting in Ottawa, 1982. ISCCP is jointly administered by the World Meteorological Organization and the International Council of Scientific Unions.

Analysis over a 15-day period in February 1979 was to be undertaken. Data from the GOES-E and TIROS-N satellites were provided under the ISCCP and distributed by F. Mosher (Wisconsin University) who also provided navigation information and attempted to remove or, at least, reduce some of the bad data.

It was chastening to find that only nine of the many cooperating scientific groups within ISCCP had managed more than a very preliminary investigation of this small time and space subset, and of these, one-third are groups providing and analysing independent data to be used later in the validation phase. All the groups presented their analysis techniques. The most popular procedure is a two-dimensional or bi-spectral analysis which determines areas of different cloud type and character on a plane defined by the visible and IR signals associated with points or small areas. An interesting clustering technique, originally developed for METEOSAT which provides a third (water vapour) channel, was described by M. Debois (CNRS, France). E. Ruprecht (University of Cologne) and P. Minnis (NASA Langley Research Center) presented results from two-dimensional schemes that provided information on cloud amount, albedo and height. The most sophisticated bi-spectral method (E. Smith, Colorado State University) allowed for dynamic calculation of surface and cloud characteristics within the routine as a function of the radiative properties of the atmosphere and also accounted for attenuation of radiation by water vapour.

Two alternative schemes were presented. One procedure (B. Rossow, GISS) inverts radiative transfer calculations to determine cloud amount, height and optical depth from the radiation fluxes emitted at the top of the atmosphere. This technique is well tested on other planets but, until now, has never been applied to the Earth. The other (J. Coakley, National Center for Atmospheric Research) uses the relationship between local mean and standard deviation in both visible and IR wavelengths. The variability of these two signals provides additional characterizing information from which cloud type and structure can be derived.

Analyses of independent data sets came from J. London (University of Colorado) and A. Henderson-Sellers (Liverpool University). London's data are derived from ship observations of cloud amount and type, whilst the Liverpool group has undertaken a detailed study of the US Air Force's three-dimensional Nephelometer. This is a hybrid data set composed of conventional and satellite information, but the remote data are derived from an independent series of polar orbiting satellites. Preliminary comparisons between the satellite and independent cloud statistics were encouraging.

An exciting recent development with potential for the validation of derived

cloud characteristics is the use of stereoscopic satellite imagery. Restricted areas of the globe can be sensed concurrently by two geostationary and a polar orbiting satellite. From these data, cloud heights and hence cloud type can be independently determined.

A major problem facing ISCCP is the enormous volume of raw satellite data that is continuously produced and transmitted to Earth. The preliminary plan calls for a 1,000-fold reduction in data volume in the first processing step. The effects on cloud determination algorithms of this data reduction process are recognized as critical. Participants were urged to study the effects of the various proposed sampling and/or averaging strategies. A further problem associated with data reduction will occur at the validation stage. Comparison between surface- and satellite-derived information is very difficult but compatibility is more likely if similar resolutions are available. Conventional data are constrained, by the geometry of the sky, to a scale of 30–50 km. The

processing strategy adopted by ISCCP must be designed to permit both later improvements in cloud retrieval procedures and validation of the cloud climatologies produced.

There are numerous problems yet to be tackled. Night-time cloud determination, for instance, is restricted to the IR channel alone. Not only is this likely to degrade the derived cloud information but it also means that considerable care must be taken to ensure compatibility between day- and night-time climatologies. The lack of illumination, the low angle of the Sun and the multiple scattering effects between cloud and snow surfaces make cloud determination in high latitudes particularly hazardous. An important cause of differences between the cloud retrieval schemes turned out to be the lack of adequate and agreed information about the surface. In particular, more detailed surveys of surface albedo and surface temperature climatologies are urgently required against which individual surface retrieval schemes can be tested. □

## Supersymmetry — fact or fiction?

from Mary K. Gaillard

PARTICLE theorists have made great progress over the past decade in unifying the apparently disparate forces of nature. Nevertheless, the arbitrary aspects of their most successful field theoretic formulations of particle interactions still leave room for dissatisfaction. One line of attack which has recently become popular is the introduction of 'supersymmetry', which relates masses and couplings of particles of different spin. This has generated an explosion of models which provide in turn a new set of candidates for the missing cosmological mass (see *J. Silk News and Views* 297, 102; 1982).

Supersymmetric theories predict many new particles, one of which should be stable, and others which may be long lived, depending on the details of the model. Candidates for the missing mass include the photino<sup>1</sup>, which is the spin  $-1/2$  superpartner of the photon; the gravitino<sup>2,3</sup>, which is the spin  $-3/2$  superpartner of the graviton; and a variety of spinless scalar particles<sup>4</sup> that can arise in supersymmetric models. In these models, the vacuum, or state of lowest energy, has a high degree of degeneracy in the classical limit; this degeneracy is removed by quantum effects. As argued by Witten<sup>5</sup>, if

the energy shift due to quantum corrections is sufficiently weak it may be possible to understand the origin of the vastly different mass scales which characterize our present picture of particle interactions — in particular, the scale  $m_W \sim 100$  GeV where the partial unification of weak and electromagnetic forces is believed to occur, and the scale  $M_X \sim 10^{15}$  GeV or more, where we anticipate their further unification with the strong forces. In such a scenario, one predicts the appearance of a light, weakly coupled scalar (Wittino) whose mass is also determined by quantum effects. Other light, weakly coupled scalars (axions) may arise in connection with additional symmetries which appear naturally in supersymmetric theories.

The upper limit on the cosmological mass density requires that quasi-stable objects such as the above be sufficiently light that their contribution to the mass density is below this limit, or sufficiently heavy that they decay or annihilate before their interactions become too slow to maintain thermal equilibrium. These constraints can be refined by requiring that any reheating of the cosmological gas from their decay or annihilation should not invalidate our present understanding of nucleosynthesis, or, if one wants to be optimistic, our 'understanding' of baryosynthesis — the process which generated the observed excess of matter

Mary K. Gaillard is in the Department of Physics, University of California, Berkeley, California 94720.



over anti-matter and which is conventionally assumed to have occurred within a tiny fraction of a second after the big bang.

The mass limits thus obtained for gravitinos and Wittinos — generally excluding a mass range between 1 KeV and  $10^4$ – $10^{11}$  GeV, under various model-dependent assumptions — are then used to exclude a range of values for the energy scale  $m_s$  above which supersymmetry would become a good approximation to nature. This is done by exploiting mass relations such as  $m_{\tilde{g}} = m_s^2/m_p$  for the gravitino mass  $m_{\tilde{g}}$ , where the Planck mass  $m_p \sim 10^{19}$  GeV is the inverse square root of Newton's constant. The Wittino mass satisfies a similar relation with  $m_p$  replaced by the unification scale,  $10^{15}$  GeV  $\leq m_{\tilde{W}} \leq m_p$ . Then, again depending on assumptions, a range of supersymmetry breaking scales  $10^3$ – $10^6$  GeV  $\leq M_s \leq 10^{11}$ – $10^{16}$  GeV appears ruled out.

These arguments suggest either that supersymmetry must be valid at an energy scale not far above the Glashow–Weinberg–Salam electroweak unification scale  $m_W$ , or that it can be valid only at a scale not very far below the strong and electroweak unification scale  $M_X$ . Models of both types exist. One should beware, however, of hasty conclusions. Supersymmetry forces us, at long last, to take quantum gravity into account, if only in adding the gravitino to our elementary particles zoo. On the other hand, simple supersymmetric theories of the type exploited for particle phenomenology do not lead to a tractable theory of quantum gravity, and the above analyses, in particular relations between 'ino' masses and the scale of supersymmetry breaking, could be altered by quantum gravity corrections, or could be invalid within the context of an ultimate theory of gravity which would become manifest at energy scales above the Planck mass.

One anticipated source of new information is from the proton decay searches that are currently underway. If supersymmetry plays an important part in the forces unifying the strong, weak and electromagnetic interactions, we can expect proton decay patterns vastly different from those predicted in more orthodox models. It will be difficult, however, to draw any firm conclusion from these data, since most models and their predictions are highly flexible. What is sorely needed is some direct experimental evidence that any superpartners of the presently known particle species do indeed exist, or, alternatively, tighter experimental constraints on the wide range of models and superparticle mass spectra for which theorists currently have a free rein for speculation.  $\square$

# The outer regions of galaxies

from Andrew Lawrence

A few decades ago, astronomers thought they knew what galaxies were and what they were made of, if not how they got there. But as so often happens, new discoveries and contemplation reveal our ignorance. So, in the last two decades, confusion has increased concerning the contents and behaviour of galaxies. The dynamics of clusters of galaxies and the rotation curves of spiral galaxies have led to the concept of 'missing mass', a dark component of galaxies, or intergalactic space, that is actually massively dominant. Deep photographs have shown some galaxies to be far larger than previously realized, sometimes with disturbed structures. Radio and X-ray studies also have revealed large quantities of gas over regions larger than the optical galaxy.

At a recent conference at the Royal Greenwich Observatory (RGO), participants examined new results and progress on 'The Outer Regions of Galaxies'. Some controversial and confusing results were presented. F. Bertola (University of Bologna) showed that the gas and the stars in some elliptical galaxies have systematically different rotation axes, and can both be different from the optical axis of the galaxy. Deep photographic plates taken with the UK Schmidt Telescope (R. Cannon, Royal Observatory, Edinburgh) show extraordinary faint features in many galaxies extending out to 100 kpc. There seems to be no consistent pattern; there are knots, wisps, blobs, streams . . . some features appear to be gaseous, some stellar. A similar story was told by P. Appleton (Jodrell Bank) concerning radio mapping of H I regions in SO galaxies. A variety of structures — haloes, rings, plumes and sometimes filaments joining members of groups of galaxies — is seen. According to R. Allen (University of Groningen), the most common H I structure for spiral galaxies is a central depression, followed by a ring, and an outer envelope. The distribution is often lopsided, however. Only some galaxies have H I envelopes, and it is not clear why.

As well as neutral gas, large quantities of very hot gas and ionized gas have been found. E.A. Valentijn (European Southern Observatory) reported 21 cm and optical CCD observations of CD galaxies in clusters with known X-ray haloes which are thought to be due to intracluster material accreting onto the giant central galaxy. Valentijn found that they got bluer towards the middle, suggesting stellar formation in the cooling accretion flow. It is the standard wisdom that spiral galaxies contain gas whereas elliptical galaxies do not. R. Fosbury (RGO) summarized observations of elliptical galaxies which do

contain gas. Many have been detected by virtue of strong nebular emission lines. The lines are of high excitation, so they do not come from normal H II regions but are probably ionized by an active nucleus. In some cases, Fosbury has mapped the velocity field of the gas with Taurus, an imaging tuneable Fabry Perot, and is able to say that the gas rotates and the stars do not. This leads to a model of a warped disk of gas cutting through an elliptical galaxy — the prototype is Centaurus A, some new Taurus observations of which were presented by P. Atherton (Imperial College, London). More conventional long-slit measurements of galaxies with enormous haloes of ionized gas were described by J. Bergeron (University of Paris). MR2251-179 has a radius of 170 kpc in the [OIII]  $\lambda$  5007 line, compared with a stellar radius of 15 kpc. A. Songaila-Cowie (RGO) discussed the deduction of properties of a hot gaseous halo around our own Galaxy by observations of absorption lines in extragalactic objects.

The puzzle of dark matter, and the existence of luminous haloes, has led to a revival of the classical area of elucidating the structure of our own Galaxy by faint star counts. J. Bahcall (Institute for Advanced Studies, Princeton) and G. Gilmore (Royal Observatory, Edinburgh) discussed their recent attempts to derive the parameters of the possible components of a spiral galaxy: disk, bulge and halo. Gilmore claimed a two-component disk, with young stars having a scale height of 100 pc, and older stars a scale height of 300 pc. On top of this, he claimed a highly flattened spheroidal 'bulge' of scale height 1.5 kpc, but no evidence of a spherical halo. CCD photometry of faint stars in Pavo (R. Gilmozzi, RGO) suggests an intermediate population II classification. S. Phillips (University of Cardiff) argued that the spheroidal component of galaxies must dominate the light of a spiral galaxy past a given radius. This may be interpreted to suggest that all spiral galaxies 'live inside' elliptical galaxies. However, M. Capaccioli (University of Padova) presented surface photometry of spiral bulges that show that they do not follow the familiar  $R^4$  intensity distribution law of elliptical galaxies, but something rather closer to  $R^{1/2}$ .

Prime evidence for the existence of massive dark haloes has been the observation of flat rotation curves. At the conference, R. Allen (University of Groningen) described observations of 18 galaxy binaries which show that their orbital masses are larger than their

1. Cabibbo, Farrar & Maniani *Phys. Lett.* **105B**, 55 (1981).
2. Pagels & Primack *Phys. Rev. Lett.* **48**, 223 (1982).
3. Weinberg *Phys. Rev. Lett.* **48**, 1303 (1982).
4. Hung & Suzuki *Preprint* University of California, Berkeley.
5. Witten *Phys. Lett.* **105B**, 267 (1981).

Andrew Lawrence is a Senior Research Fellow at the Royal Greenwich Observatory, Herstmonceux Castle, East Sussex BN27 1RP.

apparent rotation curve masses; that is that these galaxies contain substantial mass outside the last measured point of the rotation curve. Bahcall and Gilmore told the conference that such a halo in our own Galaxy could not be made of low-mass main-sequence stars without violating the observed star counts. D. Burstein (NRAO) presented rotation curves for a large sample of spiral galaxies and demonstrated that they fall into three distinct classes, which correlate roughly, but not exactly, with Hubble type. In each class, however, the rotation curves of different galaxies are very similar except for a transformation of scale.

Disturbed motions and structures in galaxies were discussed by D. Lynden-Bell (University of Cambridge), P.O. Lindblad (University of Stockholm) and R. Hook (RGO). Lynden-Bell suggested that the dwarf spheroidal galaxies Leo I, Leo II, Sculptor and Fornax, which are satellites of our Galaxy, could be part of a second stream (along with the Magellanic stream) made by a swallowed galaxy. Lindblad discussed evidence of streaming motions in our Galaxy and along the minor axis of M31. R. Hook presented Taurus observations of a ring-shaped galaxy (the Cartwheel) which showed the ring to be rotating and expanding in a regular manner. Hook suggested a model in which the system was formed by a companion galaxy passing through the larger galaxy.

A final possible constituent of galaxies is cosmic rays, and the  $\gamma$  rays produced by interactions of cosmic rays with the interstellar medium.  $\gamma$ -ray data from SAS-2 and COS-B (A.W. Wolfendale, University of Durham) indeed seem to suggest a galactic, rather than an extragalactic, origin. The flux density is greater towards the Galactic Centre, and in particular correlates with the observed density of H II regions. □

## Methylation of chloroplast DNA in *Chlamydomonas*

from T.A. Dyer

METHYLATION of DNA appears to have very important biological functions. In prokaryotes, one of its roles is to protect the organism's DNA from cleavage by sequence-specific nucleases which degrade incoming foreign DNA<sup>1</sup>, while in the nuclei of eukaryotes, changes in DNA methylation patterns may help regulate gene expression<sup>2</sup>. Although organelle DNA is generally unmethylated, some of the cytosine residues of the chloroplast DNA of the unicellular green alga *Chlamydomonas reinhardtii* are modified in this way. The methylation of chloroplast DNA is of particular interest because of the ambivalent nature of chloroplasts. They are part of eukaryotic organisms, but they have many prokaryotic features and are thought to have evolved from free-living prokaryote-like ancestors. Over the years an impressive body of evidence has been built up by Ruth Sager and her colleagues<sup>3-8</sup> to suggest that, as in prokaryotes, the methylation of chloroplast DNA in *Chlamydomonas* has a protective role. However, the recent findings of Gillham, Boynton and their associates<sup>9</sup> raise doubts about this.

The interest in methylation of the chloroplast DNA of *Chlamydomonas* currently centres on the way in which this DNA is inherited when haploid gametes of opposite mating types ( $mt^+$  and  $mt^-$ ) are mixed. Following cell fusion of the cell walls of the gametes, the chloroplasts also

fuse. But in over 90 per cent of the zygotes formed, the surviving chloroplast DNA is of the  $mt^+$  parental type only. This 'maternal'-type inheritance is found in many plants, and in higher plants it is apparently because the 'paternal' chloroplasts are excluded from the zygote<sup>10</sup>. Sager suggested that in *Chlamydomonas*, methylation of the chloroplast DNA of the  $mt^+$  parent prevents its degradation, while DNA of the  $mt^-$  parent is destroyed. Considerable circumstantial evidence supported this hypothesis. For example, the gametic  $mt^+$  chloroplast DNA contains 2 per cent 5-methylcytosine but this modified base could not be found in the  $mt^-$  gametic DNA<sup>4</sup>. The difference in level of methylation can be detected using restriction enzymes which are inhibited by methylation of bases in their recognition sequences<sup>5</sup>. Furthermore, a nuclear mutant, *mat-1*, which causes extensive gametic methylation and is tightly linked to the  $mt^-$  allele<sup>6</sup>, increases  $mt^-$  transmission of chloroplast genes to more than 50 per cent<sup>7</sup>. Finally, a DNA methyltransferase which produces 5'-methylcytosine was found only in  $mt^+$  gametes, gametes of the *mat-1* mutant and in zygotes<sup>8</sup>.

Recent findings of Gillham, Boynton and their associates<sup>9</sup> cast considerable doubt on this hypothesis. They showed that a different nuclear mutant, *me-1*, which causes extensive cytosine methylation in the chloroplast DNA of the gametes, does not increase the transmission of  $mt^-$  chloroplast DNA in crosses with wild-type  $mt^+$  stock. Indeed, the levels of methylation may be considerably

T.A. Dyer is at the Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ.

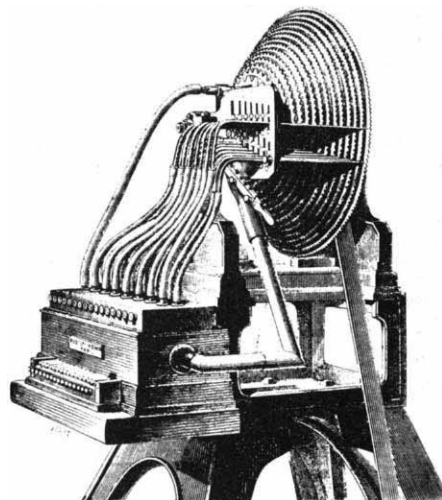


### 100 yrs ago

#### Koenig's Experiments in Acoustics

In a preceding article it was recounted how Koenig has applied the principle of the wave-siren to prove by direct experiment the influence which *phase* has upon the quality of a sound. In order to carry out these researches more fully, Koenig has constructed a very large and complete apparatus on the principle of the wave-siren. Its mode of operation will be best understood by reference to Fig. 5, taken by Dr Koenig's permission from his work, "Quelques Expériences d'Acoustique." Upon a strong stage about 4 feet high is mounted a series of 16 brass disks, cut at their edges into sinusoidal wave-forms, all fixed upon a common axis, and capable of being rotated by a band and treadle. The wave-forms cut

against the contours of these 16 disks represent a harmonic series of 16 members of decreasing amplitude, there being just 16 times as many sinuosities on the largest as on the smallest disk. Against the edge of each of these wave-disks wind can be blown by a special mouth-piece in the form of a horizontally-placed slit connected by a tube to a powerful wind-chest mounted upon the stand of the instrument. When the axis is rotated the wave-disks pass in front of the slits through which the wind is blown, and throw the issuing streams of air into vibration. Each wave-disk thus sets up a perfectly simple tone. It is clear that any desired combination can be made by opening the appropriate stops on the wind chest. In order to vary at will the *phase* in which these elementary tones are combined, a very ingenious arrangement is adopted. The brass tubes which terminate in the fifteen mouth-pieces are connected by flexible caoutchouc pipes to the wind-chest. The mouth-piece tubes are mounted upon a plate in such a way that they can slide up and down in curved slots concentric with the disks. By the aid of templates cut out in comb-fashion, and screwed to a lever handle, the mouth-pieces, or any set of them, can be displaced at will; thereby introducing any required difference of phase.



From *Nature* 26, 275, 20 July 1882.



higher in the *mt*<sup>-</sup> gametes than in the *mt*<sup>+</sup> gametes under these conditions. This methylation causes a predictable decrease in the number of target sites in the chloroplast DNA for certain methylation-sensitive restriction enzymes. Altogether, up to 40 per cent of the cytosine residues may be modified in the *me-1* mutant whereas in the wild-type *mt*<sup>+</sup> gametes only about 3 per cent of the cytosine residues are methylated. These results do not exclude the possibility that methylation at a few specific sites could be responsible for the normal maternal inheritance, but they make it seem rather unlikely.

Kuroiwa and his associates describe in this issue of *Nature* (p.481) what happens to the chloroplast DNA on mating. They were able to study the disappearance of the *mt*<sup>-</sup> chloroplast DNA with the use of epifluorescence (reflected-light fluorescence) microscopy. Using this technique, the 75 or so molecules of DNA per chloroplast, which are aggregated in about 10 'nucleoids', can be clearly distinguished in the single cup-shaped chloroplast of each gamete. Surprisingly, the *mt*<sup>-</sup> DNA disappears after fusion of the cell walls of the gametes but before the fusion of the chloroplasts. While their results do not exclude the possibility that the chloroplast DNAs of the two mating types are differently susceptible to degradation, it could be that the enzyme which destroys the *mt*<sup>-</sup> DNA is either only activated in the chloroplasts of the *mt*<sup>-</sup> gametes or can only enter the *mt*<sup>-</sup> chloroplasts due to some difference in chloroplast envelope permeability. The observation does, however, suggest that there is no competitive exclusion of the DNA from the *mt*<sup>-</sup> gametes due to a shortage of attachment sites<sup>11</sup>, as in the case with incompatible bacterial plasmids.

Kuroiwa *et al.* also show that in situations where *mt*<sup>-</sup> chloroplast DNA is destroyed, the nucleoids of the *mt*<sup>+</sup> parent coalesce after fusion of the chloroplasts and form one large nucleoid. This also takes place in the small number of 'exceptional' zygotes (less than 10 per cent) in which both the *mt*<sup>+</sup> and *mt*<sup>-</sup> DNAs persist. A similar phenomenon has been observed in mitochondria during fusion of yeast cells<sup>12</sup>, where there is a high frequency of recombination of the mitochondrial DNA.

Recombination of *Chlamydomonas* chloroplast genes also occurs at a low frequency in these exceptional zygotes. Indeed, this is the principal reason for the widespread interest in this organism as it provides a genetic method for mapping the chloroplast DNA. It is tempting to speculate that in the *Chlamydomonas* chloroplasts, recombination is made possible by the coalescence of the nucleoids containing

*mt*<sup>+</sup> and *mt*<sup>-</sup> DNAs. The role of the methylation of the DNA, however, remains perplexing, especially as it increases after zygote formation<sup>4</sup>. Could it have some role in mismatch repair, as in some prokaryotes<sup>13</sup>? Perhaps on nucleoid coalescence there is proofreading of the multiple copies of the chloroplast DNAs so that they do not diverge in sequence, and this is facilitated by the methylation. □

## Light-piping by plant tissues

from Harry Smith

THE development and application of optical fibres for long-distance communication has been justly hailed as a triumph of modern technology — but even here, the engineers will now have to admit that ingenious nature, in the form of the etiolated grass seedling, had got there first! Etiolated (that is, dark-grown) tissues, particularly the coleoptiles which sheath the first-formed leaves of grass seedlings, are remarkably sensitive to light, and, if illuminated from one side, will grow unequally on the illuminated and shaded sides, and bend towards the light source.

An article of faith in this field, traceable to the pioneer experiments of Charles Darwin and his son and published in 1880 (*The Power of Movement in Plants*), is that the site of photoperception is very close to the apex of the coleoptile, whereas the differential growth occurs towards the base. This view, based on an incomplete understanding of the optical properties of plant tissues, has been largely unchallenged during the past century, but may now need substantial reconsideration as a result of the recent demonstration by D.F. Mandoli and W.R. Briggs, of the Department of Plant Biology at the Carnegie Institution in Stanford, California, that etiolated tissues act as multiple bundles of fibre optics capable of coherent transfer of light over at least 20 mm.

In two papers<sup>1,2</sup>, Mandoli and Briggs have characterized the light-guiding properties of a range of etiolated plant tissues, by applying light, either from a light-emitting diode or a He-Ne laser, to one end of a deliberately curved tissue segment. The other end passed through an opaque holder and abutted directly onto the surface of a photomultiplier. Their photographs<sup>2</sup> show graphically that light is piped through etiolated tissues, principally via the cells themselves rather than the cell walls, to emerge at the cut end. The efficiency of axial light transmission was compared with that of other internally reflecting materials, revealing that dark-grown oat mesocotyl segments, for example, were about 10 per cent, 2.2 per cent and 0.7 per cent as efficient as a glass rod, a column of 1 per cent agar and an optical fibre, respectively, but much more

efficient than agents which only scatter light. Considerable variation in efficiency was observed between tissues, with mung bean hypocotyl hooks being only about 1.3 per cent as efficient as the oat mesocotyl.

That the tissue segments act as multiple bundles of optical fibres, not as single internally reflecting structures, was shown by varying the refractive index of the bathing medium. For a single-fibre optic, internal reflection, and thus light-piping, should be eliminated when the refractive index of the bathing medium equals that of the internal medium; although increasing the effective refractive index of the bathing medium reduced axial light transmission by up to 50–55 per cent, in no case was transmission lost. Furthermore, segments bisected longitudinally transmitted approximately half as much light as whole segments; a single fibre-optic would cease light-guiding and scatter light from the cut surface after such treatment.

A remarkable feature of light transmission in these tissues is the degree of optical coherence. An aluminium pin-hole screen directed a narrow pencil of light from the laser to one end of a tissue segment, the other end of which abutted on an array of optical fibres. The light collected by each fibre was sampled by a photo-multiplier, allowing the pattern of light transmission along the tissue to be assessed. Problems were encountered in light scattering from the distal end of the segment but, nevertheless, the degree of coherent information transfer possible substantially exceeded expectation. For example, with a segment of mung bean hypocotyl, coherent transfer occurred over at least 20 mm such that a laser beam, applied eccentrically and measured over approximately 5.3 cell diameters at the point of interception with the tissue, did not spread laterally to the other side of the tissue, 20 cells away.

These unexpected optical properties of plant tissues clearly have important implications for both phototropism and photomorphogenesis. With certain

Harry Smith is Head of the Department of Botany at the University of Leicester, Leicester LE1 7RH.

1. McClelland, M. *Nucleic Acids Res.* **9**, 5859 (1981).
2. Felsenfeld, G. & McGhee, J.D. *Nature* **296**, 602 (1982).
3. Sager, R. & Lane, D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2410 (1972).
4. Burton, W.G., Grabow, C.T. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1390 (1979).
5. Royer, H.-D. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5794 (1979).
6. Sager, R., Grabow, C. & Sano, H. *Cell* **24**, 41 (1981).
7. Sager, R. & Ramanis, Z. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4698 (1974).
8. Sano, H. & Sager, R. *Eur. J. Biochem.* **105**, 471 (1980).
9. Bolen, P.L., Grant, D.M., Swinton, D., Boynton, J.E. & Gillham, N.W. *Cell* **28**, 335 (1982).
10. Kirk, J.T.O. & Tilney-Basset, R.A.E. *The Plastids. Their Chemistry, Structure, Growth and Inheritance* (Elsevier, Amsterdam, 1978).
11. Gillham, N.W., Boynton, J.E. & Lee, R.W. *Genetics* **78**, 439 (1974).
12. Sando, N., Miyakawa, I., Nishibayashi, S. & Kuroiwa, T. *J. gen. appl. Microbiol.* **27**, 311 (1981).
13. Glickman, B.W., van den Essen, P. & Radman, M. *Molec. gen. Gennet.* **163**, 307 (1978).





partly explained by analytical difficulties. The polymer cannot be specifically radioactively labelled, except in cells made permeable to NAD. Analysis of poly(ADP-ribose) formation has therefore depended on awkward procedures involving degradation of the polymer to phosphoribosyl-AMP (by cleavage of the phosphate-phosphate bonds with snake venom phosphodiesterase) and quantification of the latter compound, or on access to antibodies recognizing poly(ADP-ribose). Effective monoclonal antibodies are not, however, available. Fortunately, a series of nicotinamide analogues (for example, 3-amino-benzamide) act as specific inhibitors of poly(ADP-ribose) polymerase, and several important studies on the role of poly(ADP-ribose) rely heavily on the use of such compounds.

A particularly interesting finding<sup>11</sup> is that treatment of mammalian cells with these inhibitors in the presence of bromodeoxyuridine causes a substantial increase in sister chromatid exchange. The recombination observed is largely due to the occurrence of bromodeoxyuridine in DNA

(which is necessary in order to visualize chromatid exchanges), because the incorporation of this hydrolytically labile residue leads to an increased background of spontaneously occurring DNA chain scissions<sup>12</sup>. These strand interruptions would facilitate recombination events as long as they remain open, but they could also serve to induce poly(ADP-ribose) formation and thereby trigger their own repair. Indeed, the extent to which the detour repair pathway has been activated may be of general importance in assessments of the relative sensitivities of mammalian cells under different conditions to DNA-damaging agents. □

1. Hayaishi, O. & Ueda, K. *A. Rev. Biochem.* **46**, 95 (1977).
2. Shall, S. *Adv. Radiat. Biol.* (in the press).
3. Taniguchi, T. *et al. J. biol. Chem.* **257**, 4027 (1982).
4. James, M. R. & Lehmann, A. R. *Biochemistry* (in the press).
5. Creissen, D. & Shall, S. *Nature* **296**, 271 (1982).
6. Soderhall, S. *Nature* **260**, 640 (1976).
7. Suzuki, K. *et al. FEBS Lett.* **119**, 150 (1980).
8. Berger, N. A. & Sikorski, G. W. *Biochemistry* **20**, 3610 (1981); Poirier, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3423 (1982).
9. Paulson, J. R. & Laemmli, U. K. *Cell* **12**, 817 (1977).
10. Jump, D. B. & Smulson, M. *Biochemistry* **19**, 1024 (1980); Ogata, N. *et al. J. biol. Chem.* **256**, 4135 (1981).
11. Oikawa, A. *et al. Biochem. biophys. Res. Commun.* **97**, 1311 (1980).
12. Natarajan, A. T. *et al. Mutat. Res.* **84**, 125 (1981).

## How not to measure inclusive fitness

from Alan Grafen

THE introduction of the concept of inclusive fitness by Hamilton<sup>1</sup> in 1964 led to a revolution in thinking about the evolution of social behaviour, by making clear how selection can act on a gene through its effects on its bearer's relatives. This led to a better understanding of the evolution of insect societies, and by 1977 it had been stated as the 'central theorem' of behavioural biology<sup>2</sup> that 'animals maximize their inclusive fitness'. Detailed studies of populations of individually marked animals are now being performed, and it is essential that when the data collected are used to calculate inclusive fitness the correct measure is employed.

Several recent studies<sup>3,4</sup> give cause for considerable concern that erroneous measures that look like, but are in fact different from, inclusive fitness are becoming prevalent. This problem can also be seen by surveying fourteen recently published books<sup>5-18</sup>, including the most used animal behaviour textbooks. All use the term inclusive fitness, but of the ten that do define it, none does so correctly. In a loosely argued discussion, the misdefinition may not lead to serious error. For modellers and for field workers with data, on the other hand, the precise definition matters a great deal.

One erroneous measure, which can be called 'simple weighted sum' (SWS), may

be defined by borrowing Barash's erroneous definition of inclusive fitness<sup>9</sup>:

"the sum of individual fitness (reproductive success) and the reproductive success of an individual's relatives, with each relative devalued in proportion as it is more distantly related"

As we never know about all of an animal's relatives, SWS itself has never been measured. Versions of SWS where only certain relatives are included have been used by Bygott *et al.*<sup>3</sup> and McGregor *et al.*<sup>4</sup>. Hamilton<sup>1</sup> described inclusive fitness as

"the animal's production of adult offspring . . . stripped of all components . . . due to the individual's social environment, leaving the fitness he would express if not exposed to any of the harms or benefits of that environment, . . . and augmented by certain fractions of the quantities of the harm and benefit the individual himself causes to the fitnesses of his neighbours. The fractions in question are simply the coefficients of relationship . . ."

Two major differences emerge. First, SWS counts all the animal's own offspring while inclusive fitness counts the number the

animal would have had in the absence of help and hindrance from other animals. Second, SWS counts all relatives' offspring. Inclusive fitness counts only those that owe their existence to the animal's help, and counts negatively those relatives' offspring that would have existed but for the hindrance of the animal.

The contribution to SWS from an animal's own offspring may be greater or smaller than that contribution to inclusive fitness. The contribution from relatives' offspring must be positive to SWS, but may be positive or negative to inclusive fitness.

In many cases it will be extremely difficult to assess how many offspring an animal would have had but for the intervention of another animal, yet this must be done to measure inclusive fitness. There is, however, a simpler way to include kinship effects in a measure of reproductive success — simply count the number of the animal's offspring. Such a measure inevitably predicts the course of evolution for, if animals bearing an allele have more offspring than non-bearers, then the allele will spread. It is perhaps not commonly appreciated that this measure also includes kinship effects.

If I have an allele for helping sibs, then my sibs gain through my help. But since they are more likely to share my alleles than is the population as a whole, they will help me and so their contribution is included in my number of offspring. Even in a case where I help but am not helped in return (suppose in a clutch of young birds, the first hatching helps the second break out of its shell but the second cannot return this kind of help), kinship effects are still included when we average over all bearers of the helping allele, including first and second hatchlings.

If number of offspring is a measure of reproductive success that encompasses kinship effects, why did Hamilton invent inclusive fitness and why is the central theorem phrased in terms of it? The importance of Hamilton's measure is that, in calculating whether a hypothetical behavioural trait will spread, it is equivalent to the number of offspring measure but is often very much simpler to use and much more revealing. In particular it leads immediately to Hamilton's rule ( $B/C > 1/r$ ), where  $B$  is the benefit to the recipient,  $C$  is the cost to the actor (both measured in number of offspring), and  $r$  is the coefficient of relatedness, which expresses clearly the role of relatedness. The two measures are equivalent in that they always make the same prediction about the direction of selection of an allele, although they will not usually be numerically equal. It is this equivalence that Hamilton<sup>1</sup> proved in 1964 to justify inclusive fitness.

Maynard Smith<sup>19</sup> provides a well worked out example to illustrate the computational and conceptual advantages of inclusive fitness in modelling. He considers a diploid population with two alleles  $A$  and  $a$  at one locus.  $AA$  and  $Aa$  genotypes are

Alan Grafen is in the Animal Behaviour Research Group, Department of Zoology, Oxford OX1 3PS and a Research Lecturer of Christ Church College, Oxford. Marian Dawkins, Mark Ridley and Richard Dawkins are thanked for their comments on an earlier draft of the manuscript.

altruistic at cost  $C$  to the donor and benefit  $B$  to the recipient. Two separate cases are considered: in the first, the recipients are all half-sibs while in the second they are all double first cousins. What is the condition for  $A$  to spread?

The inclusive fitness method gives the answer directly, through Hamilton's rule. In both cases  $r = 1/4$ , and so in both cases  $A$  will spread if  $B/C > 4$ .

The standard population genetics method, counting the expected number of offspring of animals with each genotype, leads to one set of three equations for each case, and the equations are different in the two cases. The condition for  $A$  to spread can be derived from either set of equations, and turns out to be independent of the frequency of  $A$  (though all three equations involve this variable), and in fact to be  $B/C > 4$ . The important points are that in this method each different kind of relative has to be dealt with differently, obscuring the single important parameter  $r$ , and that considerable mathematical skill is needed to derive correctly the set of equations and to deduce correctly from them. These disadvantages are overcome by Hamilton's rule at the cost of making some simplifying assumptions that are likely to be true enough for data analysis and rough modelling<sup>20,21</sup>.

There are thus two valid measures of reproductive success. Number of offspring will usually be more suitable in field studies where the mean numbers can easily be compared between groups of interest. Inclusive fitness will usually be more suitable for modelling, when the hypothetical trait in question is defined in terms of differences in numbers of offspring. Inclusive fitness can be used in field studies when we have a particular decision in mind that animals can make, when the logic of this decision guides us into calculating entirely in terms of differences in numbers of offspring, and where by experiment or natural variation we can measure those differences. For an example of this see Emlen's discussion<sup>22</sup> of helping at the nest in birds.

The errors that SWS can lead to as a measure of reproductive success can be documented in two recent cases where a restricted version of SWS has been calculated from field data. Bygott *et al.*<sup>3</sup> were assessing the evolutionary advantage a male lion gains through belonging to a large group of males — larger groups keep control of a pride of females longer and may also control more prides. Bygott *et al.* presented two different measures of reproductive success as a function of male group size. The first was number of offspring in a lifetime, and was higher in groups of three and more than in groups of one or two. The second was calculated as: sum of (a) the animal's offspring and (b) 0.22 times the total number of his fellow group members' offspring (0.22 is an estimate of the average coefficient of relatedness between two males in a group). Not surprisingly this

restricted version of SWS shows a dramatic increase as group size increases. The authors used this result to argue that the advantage of larger groups was so great that relaxing their assumption of equal matings by males in a group would not bring the advantage into doubt, even for the least successful members of a large group.

McGregor *et al.*<sup>4</sup> studied the evolutionary significance of repertory size in the song of male great tits. They showed that number of offspring increased with repertory size and then went on to calculate a restricted version of SWS, adding to number of offspring the following relatives appropriately devalued: grandoffspring, great-grandoffspring, nieces, nephews, great-nieces and great-nephews. Using this measure, the population mean repertory size was close to the size that maximized reproductive success. McGregor *et al.* concluded that there were indirect effects at work that favoured birds with small repertory size, compensating them for their small number of offspring; and that repertory size is maintained near its optimum by natural selection. Since these conclusions contradict the finding that a valid measure of reproductive success, number of offspring, increases with repertory size, they cannot be correct. We can attribute the error to use of SWS.

Two theoretical points can be used to show up the flaws in SWS. First, in a simple and standard populations genetics model, SWS cannot be calculated because it is infinitely large. In a diploid panmictic lifetime-monogamous infinite population

with no selection, and with discrete generations, consider only those animals in my generation that are related to me through a common ancestor no more than  $n$  generations ago. The contribution of those animals to my SWS is  $1 + (n/2)$ . As we count more and more distant relatives this becomes indefinitely large, as all components of SWS are positive.

Second, even when SWS is restricted to animals whose reproduction can be influenced by an individual, it gives erroneous results. Consider an infinite panmictic monogamous diploid species with discrete generations and sibships of size  $s$ . Let there be a rare dominant allele  $A$  that causes its bearers to decrease its own number of offspring by  $c$  while increasing its sibs' number of offspring by a total of  $b$  distributed equally, and suppose these costs and benefits combine additively. The condition for  $A$  to spread initially is

$$\frac{1}{2} b - c > 0$$

This follows from both the inclusive fitness method and the full population genetics model. I have computed the condition for  $A$  to spread using the following restricted version of SWS: the sum of (a) the individual's offspring plus (b) 0.5 times the total number of offspring of the other sibs in the individual's sibship. That condition is

$$((s+2)/(s+3))b - c > 0$$

This condition is different from that above and gives an erroneous measure that is systematically too favourable to altruistic alleles.

Using the same example we can calculate the condition for  $A$  to spread according to the other common misdefinition of inclusive fitness, given, for example, by Wilson<sup>14</sup>; it is the sum of (a) the individual's own number of offspring and (b) the effects of the individual on its relatives' number of offspring weighted by the degrees of relatedness. This measure gives the condition for  $A$  to spread as

$$b - c > 0$$

This measure is also systematically too favourable to the altruistic allele.

Of the ten books that were mentioned earlier as misdefining inclusive fitness, five<sup>9-13</sup> define it as SWS, as have (astonishingly) Maynard Smith<sup>23</sup> and (regrettably) myself<sup>24</sup>. The remaining five books<sup>14-18</sup> follow Wilson<sup>14</sup> in giving the definition discussed in the previous paragraph.

In conclusion, not all measures of reproductive success are equal. SWS and restrictions of it are erroneous. Two correct and convenient measures are inclusive fitness and number of offspring. Unfortunately, many modern texts that extol the virtues of inclusive fitness either fail to define it or define it wrongly. Analysers of expensive or extensive data should make very sure that they are using a valid measure of reproductive success. □

1. Hamilton, W. D. J. *theor. Biol.* 7, 1 (1964).
2. Barash, D. P. *Sociobiology and Behavior*, 63 (Elsevier, New York, 1977).
3. Bygott, J. D. *et al. Nature* 282, 839 (1979).
4. McGregor, P. K. *et al. Am. Nat.* 118, 149 (1981).
5. Brown, J. L. *The Evolution of Behavior*, 204 (Norton, New York, 1975).
6. Manning, A. *An Introduction to Animal Behaviour* 3rd edn, 211 (Arnold, London, 1979).
7. Barnett, S. A. *Modern Ethology*, 592 (Oxford University Press, 1981).
8. Broom, D. M. *Biology of Behaviour*, 9 (Cambridge University Press, 1981).
9. Barash, D. P. in *Sociobiology: Beyond Nature/Nurture?* (eds Barlow, G. W. & Silverberg, J.) 212 (Westview, Boulder, 1980).
10. Daly, M. & Wilson, M. *Sex, Evolution and Behavior*, 376 (Duxbury, Massachusetts, 1978).
11. Keeton, W. T. *Biological Science* 3rd edn, 787 (Norton, New York, 1980).
12. Krebs, J. R. & Davies, N. B. *An Introduction to Behavioural Ecology*, 20 (Blackwell, Oxford, 1981).
13. Hinde, R. A. *Ethology*, 147 (Oxford University Press, 1982).
14. Wilson, E. O. *Sociobiology*, 586 (Belknap, Cambridge, Massachusetts, 1975).
15. Dewsbury, D. A. *Comparative Animal Behaviour* 44, (McGraw-Hill, New York, 1978).
16. Alcock, J. *Animal Behavior* 2nd edn, 9 (Sinauer, Sunderland, Massachusetts, 1979).
17. Wittenberger, J. F. *Animal Social Behavior*, 62 (Duxbury, Boston, 1981).
18. Dawkins, R. *The Extended Phenotype*, 186 (Freeman, Oxford, 1982).
19. Maynard Smith, J. in *Current Problems in Sociobiology* (ed. King's College Sociobiology Group) (Cambridge University Press, in the press).
20. Charlesworth, B. in *Evolution of Social Behavior* (ed. Markl, H.) 11 (Dahlem Konferenzen, Berlin, 1980).
21. Cavalli-Sforza, L. & Feldman, M. W. *Theor. Population Biol.* 14, 268 (1978).
22. Emlen, S. T. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 245 (Blackwell, Oxford, 1978).
23. Maynard Smith, J. A. *Rev. Ecol. Systematics*, 51 (1978).
24. Grafen, A. *Anim. Behav.* 27, 907 (1979).



## ARTICLES

## QSO Lyman limit absorption

David Tytler

Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

*The redshift distribution of QSO absorption systems which are optically thick in the Lyman continuum matches that of a non-evolving population of absorbers in a standard Friedmann cosmological model over the redshift range 0.4–3.5. The density of absorbers per unit velocity in the QSO rest frame is roughly constant for  $-0.01 \leq v/c \leq 0.2$  and shows a rapid increase with QSO emission redshift, in accord with an 'intervening' origin for the absorbers but contrary to expectation were the material intrinsic to the QSOs.*

THE material causing absorption in QSO spectra is observed at a wide range of redshifts. There are two possible origins for the absorbing material. In the intrinsic hypothesis the material is physically associated with the QSO<sup>1</sup>. It may have been ejected at velocities which approach the speed of light, but will be cosmologically close ( $\leq$  few Mpc) to the QSO. In contrast in the intervening hypothesis the material is cosmologically distant from the QSO, coincidentally lying along the line of sight.

QSO absorption line systems may be classified into various types on the basis of their observed spectroscopic properties<sup>2</sup>. Sargent *et al.*<sup>3</sup> have studied the distribution of 187 Ly $\alpha$  systems. They find the same density of systems per unit  $z_{\text{abs}}$  along the lines of sight towards five different QSOs suggesting that these systems are intervening; unless an intrinsic model can be found which gives constant numbers of absorption systems for all QSOs. Similar results apply to narrow metal line systems<sup>4–6</sup>, although these samples are still small.

QSOs at cosmological distances are bound to show some intervening absorption—especially from extended galactic coronae and haloes<sup>7,8</sup>. The outstanding question is: how many, if any<sup>9,10</sup>, of the narrow line systems are intrinsic? In particular do the intrinsic broad, trough-like systems seen in  $\sim 5\%$  of QSOs break up and evolve into intrinsic narrow line systems<sup>5,11</sup>?

This article presents a statistical analysis of the distribution of a new homogeneous sample of narrow line absorption systems. Evidence is found in support of an intervening origin for these systems.

## Lyman limit systems

Lyman limit systems (LLS) are the subcategory of the narrow line absorption systems which are optically thick to photons capable of ionizing hydrogen. They deserve special attention for the following reasons:

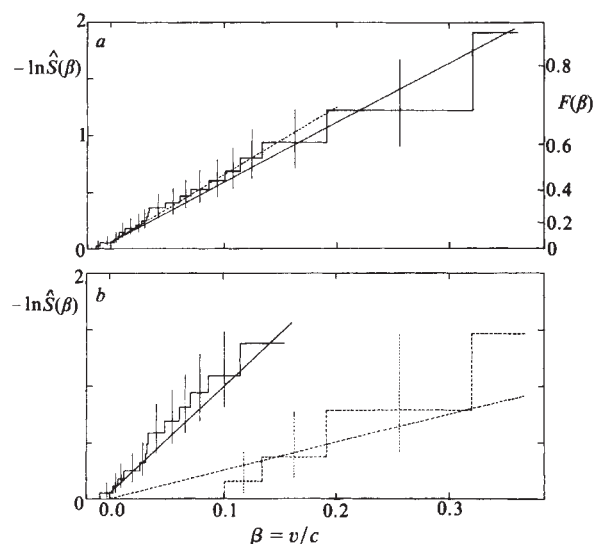
- (1) They are extremely easy to detect<sup>12</sup>. The break in a spectrum near the Lyman limit rest wavelength of 912 Å can be seen in low resolution spectra of low signal-to-noise ratio: a striking contrast to the exceptionally high quality spectra needed to identify narrow absorption lines.
- (2) Ease of detection means that we can combine optical and ultraviolet (IUE) data to obtain a homogeneous sample covering the unprecedentedly large redshift range 0.4–3.5.
- (3) LLS are included in the sample on the criterion that their column densities in H I should be above a set minimum. An H I column density is a more useful marker than the C IV or Mg II absorption line equivalent widths used to delimit samples of metal line systems<sup>5,6</sup>. In addition an LLS sample will detect systems irrespective of the velocity dispersion of the absorbing gas. In contrast, samples of systems selected through the observation of lines are biased against the detection of systems with saturated lines of low velocity dispersion, because of the low equivalent widths of such lines.
- (4) Finally note that the typical weak line absorption system is not optically thick in the Lyman continuum. The LLS are physically distinct and may have different origins.

Radiation of wavelength  $\lambda < 912$  Å with an initial intensity  $I_0$  will, on passing through a cloud of neutral hydrogen column density  $N$  (H I)  $\text{cm}^{-2}$ , emerge with an attenuated intensity  $I = I_0 \exp(-\tau_\lambda)$ , where the optical depth to Lyman continuum radiation is  $\tau_\lambda = 6.3 \times 10^{-18} N$  (H I)  $(\lambda/912)^3$  for  $\lambda \leq 912$  Å.

Table 1a contains the Lyman limit system sample. There are 36 QSOs and 20 LLS. LLS are included in the sample if they satisfy two criteria: (1)  $\tau > 1$  or  $N(\text{H I}) > 1.6 \times 10^{17} \text{ cm}^{-2}$ , corresponding to  $I < 0.368 I_0$ . (2) In the exceptional case where more than one LLS with  $\tau > 1$  has been seen in a single line of sight (for example, QSO OH471 = 0642 + 44) we include only the LLS with the highest redshift, for statistical reasons.

QSOs contribute to the sample if: (1) observations are sufficient to have revealed any such LLS; (2) I omit QSOs showing broad or trough absorption because the spectra are too complex to be certain of the presence of a weak  $\tau \approx 1$  LLS; (3) BL Lac objects without  $z_{\text{em}}$  values are omitted to simplify the analysis.

Column 2 (Table 1a) lists  $\lambda_{\text{min}}$ , the minimum wavelength at which a Lyman limit break with  $\tau > 1$  could have been detected



**Fig. 1** Cumulative Hazard plots. The negative natural logarithm of the Kaplan-Meier estimator of the survival function<sup>16</sup>  $\hat{S}$  is shown as a function of velocity  $v = c\beta$  in the QSO frame of reference, for the 20 LLS from the 36 QSOs in Table 1a. Vertical jumps occur at the positions of the LLS. The height of a jump depends on the number of QSOs towards which an LLS could have been seen, at that  $\beta$  value. *a*, The solid straight line is a maximum likelihood fit to the exponential model of equations (7) and (8), given by  $-\ln \hat{S}(\beta) = \hat{\lambda}\beta$  for  $\hat{\lambda} = 5.3$  and  $\sigma(\hat{\lambda}) = 1.2$ , obtained using equations (9) and (10). The dashed line is for  $\hat{\lambda} = 6.0$ ,  $\sigma(\hat{\lambda}) = 1.4$  obtained with *a posteriori* censoring of all data at  $\beta = 0.2$ .  $F(\beta) = 1 - \hat{S}(\beta)$  is the cumulative distribution function, giving the probability that an LLS is observed at some velocity  $< c\beta$ .  $F(\beta)$  is LLS QSO 'covering factor'<sup>12</sup> expressed as a function of  $\beta$ . *b*, QSO redshift division. Solid lines are for the 18 QSOs with  $z_{\text{em}} \geq 3$ . The straight solid line has the gradient  $\hat{\lambda} = 9.3$ ,  $\sigma(\hat{\lambda}) = 2.7$ . Dashed lines are for the seven QSOs with  $z_{\text{em}} \leq 2.2$ . The straight dashed line has  $\hat{\lambda} = 2.5$ ,  $\sigma(\hat{\lambda}) = 1.2$ .

unambiguously. I set  $\lambda_{\min}$  to the minimum observed wavelength plus 100 Å for low resolution (FWHM  $\Delta\lambda \geq 5$  Å) spectra, or plus  $\sim 30$  Å for high resolution or signal-to-noise ratio spectra. For the seven IUE spectra we cautiously choose  $\lambda_{\min} = 1,260$  Å. At  $\lambda < \lambda_{\min}$  a LLS might have been confused with the strongest of Ly $\alpha$  lines which occasionally have FWHM up to 80 Å in the observed frame<sup>13,14</sup>.

Column 3 (Table 1a) gives  $\lambda_b$  the observed wavelength of the Lyman limit break, taken at  $I(\lambda_b) = \frac{1}{2}[I_0 + I(\lambda < \lambda_b)]$ . Whenever possible  $\lambda_b$  has been remeasured to conform with this definition of the half power point.

Redshifts  $z_{\text{LLS}}$  for the LLS have been determined as follows. Column 3 of Table 1b lists  $z_{\text{abs}}$  (determined from lines alone, without reference to Lyman limit breaks) for those systems which can be regarded as probably or certainly associated with Lyman limit breaks. Systems consisting only of single Ly $\alpha$  lines are excluded. Ly $\alpha$  lines with rest frame equivalent width  $W > 0.32$  Å are extremely common<sup>3</sup>, yet the Ly $\alpha$  in a LLS with  $\tau = 1$  may even have  $W < 0.32$  Å, if the velocity dispersion in the gas has  $b < 8$  km s<sup>-1</sup>, where the velocity distribution is of form  $\exp(-v^2/b^2)$  by assumption.

The effective rest frame wavelengths  $\lambda_{\text{rb}}$  corresponding to

**Table 1** The Lyman limit systems

**a** The statistical sample

| QSO         | $\lambda_{\min}$ (Å) | $\lambda_b$ (Å) | $z_{\min}$ | $z_{\text{LLS}}$ | $z_{\text{em}}$ | $m_v$ | $-M_v$ | $\Delta\lambda$ (Å) | Refs       |
|-------------|----------------------|-----------------|------------|------------------|-----------------|-------|--------|---------------------|------------|
| 0000-398    | 3,300                | 3,488           | 2.595      | 2.800            | 2.83            | 18.8  | 27.4   | 30                  | 12         |
| 0002-422    | 3,300                | —               | 2.595      | —                | 2.763           | 17.4  | 28.7   | 30                  | 12, 25     |
| 0042.8-2657 | 3,300                | —               | 2.595      | —                | 2.90            | 19    | 27.3   | 8                   | 12         |
| 0049-393    | 3,300                | 3,445           | 2.595      | 2.753            | 2.85            | 18.0  | 28.2   | 30                  | 12         |
| 0100+130    | 3,170                | —               | 2.453      | —                | 2.690           | 16.6  | 29.5   | 2                   | 3, 13      |
| 0101.5-3025 | 3,300                | 3,565           | 2.595      | 2.883            | 3.17            | 19    | 27.6   | 8                   | 12         |
| 0103.9-2913 | 3,300                | —               | 2.595      | —                | 2.80            | ?     | ?      | 8                   | 12         |
| 0130-403    | 3,300                | —               | 2.595      | —                | 3.02            | 17.4  | 29.0   | 30                  | 12, 26*    |
| 0138-381    | 3,300                | —               | 2.595      | —                | 2.87            | 17.6  | 28.6   | 30                  | 12†        |
| 0140-306    | 3,300                | 3,665           | 2.595      | 2.992            | 3.13            | 18.5  | 28.0   | 8                   | 12         |
| 0143-015    | 3,300                | —               | 2.595      | —                | 3.13            | 18.8  | 27.7   | 30                  | 12         |
| 0143-010    | 3,300                | 3,512           | 2.595      | 2.82             | 3.17            | 19.0  | 27.6   | 30                  | 12         |
| 0207-003    | 3,300                | —               | 2.595      | —                | 2.85            | 17.7  | 28.5   | 30                  | 12‡        |
| 0207-398    | 3,300                | —               | 2.595      | —                | 2.81            | 17.5  | 28.7   | 30                  | 12         |
| 0244.4-3017 | 3,300?               | 3,741           | 2.595?     | 3.088            | 3.103           | ?     | ?      | 10                  | 27§        |
| 0347-383    | 3,300                | 3,687           | 2.595      | 3.02             | 3.22            | 17.3  | 29.3   | 30                  | 12, 26     |
| 0351-390    | 3,300                | —               | 2.595      | —                | 3.00            | 17.9  | 28.5   | 30                  | 12, 26     |
| 0405-123    | 1,260                | —               | 0.373      | —                | 0.574           | 15.0  | 26.5   | 10                  | 29, 29     |
| 0420-388    | 3,390                | 3,758           | 2.693      | 3.09             | 3.12            | 16.9  | 29.6   | 2                   | 12, 44     |
| 0528-250    | 3,300                | $\sim 3,520$    | 2.595      | 2.812            | 2.765           | 17.7  | 28.4   | 2                   | 14¶        |
| 0537-286    | 3,350                | 3,664           | 2.649      | 2.976            | 3.11            | 20.0  | 26.5   | 10                  | 30         |
| 0642+44     | 3,730                | 3,994           | 3.063      | 3.343            | 3.40            | 18.5  | 28.2   | 30                  | 31#        |
| 0805+05     | 3,400                | —               | 2.704      | —                | 2.877           | 18.2  | 28.1   | 2, 80               | 32, 33**   |
| 0938+119    | 3,500                | 3,617           | 2.813      | 2.940            | 3.19            | 19.0  | 27.6   | 20                  | 34         |
| 0955+326    | 1,260                | —               | 0.373      | —                | 0.533           | 15.8  | 25.6   | 10                  | 35         |
| 1101-264    | 1,260                | 2,600           | 0.373      | 1.8387           | 2.143           | 16.0  | 29.3   | 8                   | 37, 38     |
| 1115+080    | 1,260                | —               | 0.373      | —                | 1.718           | 18.7  | 26.0   | 10                  | 28, 39††   |
| 1225+310    | 1,260                | 2,565           | 0.373      | 1.795            | 2.200           | 15.9  | 29.5   | 8                   | 3, 40      |
| 1247+268    | 1,260                | $\sim 2,000$    | 0.373      | 1.179            | 2.038           | 15.8  | 29.4   | 10                  | 28         |
| 1317+277    | 1,260                | 1,529           | 0.373      | 0.666            | 1.022           | 15.3  | 27.7   | 8                   | 35         |
| 1402+044    | 3,740                | 3,890           | 3.074      | 3.237            | 3.20            | 18.5  | 28.1   | 8                   | 41         |
| 1442+101    | 3,230                | —               | 2.518      | —                | 3.53            | 17.8  | 29.0   | 7                   | 42         |
| 2126-158    | 3,300?               | $\sim 3,500$    | 2.595?     | 2.813            | 3.280           | 17.3  | 29.3   | 2                   | 3, 30, 43§ |
| 2204-408    | 3,300                | —               | 2.595      | —                | 3.17            | 17.5  | 29.1   | 30                  | 12‡‡       |
| 2228.2-403  | 3,450                | 3,690           | 2.758      | 3.020            | 3.15            | 19.6  | 26.9   | 40                  | 26         |
| 2348-011    | 3,300                | 3,591           | 2.595      | 2.912            | 3.02            | 19.5  | 26.9   | 30                  | 12         |

**b** Systems with well established  $z_{\text{abs}}$

| QSO or BL Lac | $\lambda_b$ (Å) | $z_{\text{abs}}$    | $\lambda_{\text{rb}}$ (Å) | Refs     |
|---------------|-----------------|---------------------|---------------------------|----------|
| 0244.4-3017   | 3,741 $\pm$ 3   | 3.088 $\pm$ 0.001   | 915.1 $\pm$ 0.8           | 27       |
| 0420-388      | 3,758 $\pm$ 1   | 3.090 $\pm$ 0.001   | 918.8 $\pm$ 0.3           | 12, 44   |
| 0528-250      | 3,520 $\pm$ 5   | 2.812 $\pm$ 0.001   | 923.4 $\pm$ 1.3           | 14       |
| 0537-286      | 3,664 $\pm$ 5   | 2.976 $\pm$ 0.001   | 921.5 $\pm$ 1.3           | 30       |
| 0642+44       | 3,994 $\pm$ 10  | 3.343 $\pm$ 0.002   | 919.6 $\pm$ 2.3           | 31#      |
| 0642+44       | 3,778 $\pm$ 10  | 3.122 $\pm$ 0.002   | 916.5 $\pm$ 2.5           | 31#      |
| 0735+178      | 1,308 $\pm$ 3   | 0.4243 $\pm$ 0.0003 | 918.3 $\pm$ 2.1           | 45, 46§§ |
| 1101-264      | 2,600 $\pm$ 8   | 1.8387 $\pm$ 0.0003 | 915.9 $\pm$ 2.8           | 37       |
| 1225+310      | 2,565 $\pm$ 10  | 1.795 $\pm$ 0.003   | 917.7 $\pm$ 3.7           | 25, 40   |
|               |                 |                     | Mean                      | 918.5    |

Magnitudes  $m_v$  are from ref. 47 except for: 0042.8-2657 and 0101.5-3025 (ref. 12), 0405-123 (refs 2, 9), 1115+080 (ref. 39,  $m_v$  corrected for gravitational amplification) and 1247+268 (ref. 28).  $M_v$  is the  $q_0=0$  absolute magnitude for a Hubble constant of 100 km s<sup>-1</sup> Mpc<sup>-1</sup>, with  $K$  corrections from ref. 48.  $\Delta\lambda$  is the instrumental resolution in the observed frame near  $\lambda_b$ . Three QSOs which have been observed in the Lyman limit wavelength range which were excluded from the sample: 0324-407 (has trough/broad line absorption<sup>12</sup>), 1704+608 (any LLS would have  $\lambda_b < 1,260$  Å), 2227.6-3928 shows a break of  $I/I_0 \sim 0.3$  at  $\lambda_b \sim 4,420$  Å. This is unlikely to be a Lyman limit edge<sup>12</sup> because it would have  $z_{\text{LLS}} = 3.81$ ; implausibly larger than  $z_{\text{em}} = 3.45$ . Further observations are required.

\* 0130-403: refs 12 and 26 both show LLS with  $\lambda_b = 3,500$  Å and  $I/I_0 \sim 0.5$ . Excluded because  $\tau < 1$ .

† 0138-381: possible LLS<sup>12</sup> excluded because either  $I/I_0 \sim 0.5$  or  $\lambda_b < 3,300$  Å.

‡ 0207-003: definite LLS<sup>12</sup> with  $I/I_0 \leq 0.04$ , but excluded because  $\lambda_b < 3,300$  Å.

§ 0244.4-3017 and 2126-158:  $\lambda_{\min}$  and  $z_{\min}$  are guesses. (They are only used in combination with all other  $z_{\min}$  values, to estimate the errors on parameters.)

|| 0351-390: possible LLS<sup>12</sup> excluded because  $\lambda_b < 3,300$  Å.

¶ 0528-250: signal-to-noise ratio very low near  $\lambda_b$ .

# 0642+44 shows at least two and possibly four LLS<sup>31</sup>. The LLS with  $\lambda_b = 3,994$  Å has  $I/I_0 \sim 0.36$  just giving sufficient  $\tau$  for inclusion in the sample. A second LLS at  $\lambda_b = 3,778$  has  $I/I_0 \leq 0.1$  but is excluded from sample since the former LLS has a higher  $z_{\text{LLS}}$  and is included.

\*\* 0805+05: spectrophotometry with 80 Å resolution by Oke<sup>32</sup> shows a possible LLS with  $I/I_0 \sim 0.4$  in range 3,410-3,480 Å, and probably near 3,445 Å. However, the 2 Å resolution spectrum of Chen *et al.*<sup>33</sup> fails to show any sudden breaks. I tentatively conclude that there are no LLS with  $\tau > 1$  ( $I/I_0 < 0.37$ ) and  $\lambda_b > 3,400$  Å.

†† 1115+080: magnitudes corrected for gravitational lense amplification<sup>39</sup>.

‡‡ 2204-408 has been detected in the UV 1,200-1,950 Å by Wilson *et al.*<sup>49</sup>. This observation is excluded because the signal-to-noise ratio and spectral resolution are so low that any absence of flux would not necessarily have been interpreted as being due to an LLS.

§§ 0735+178 is a BL Lac object, of unknown  $z_{\text{em}}$ , excluded from sample. ||| Also A. Wright, personal communication.



the observed frame  $\lambda_b$  wavelengths are given by:  $\lambda_{rb} = \lambda_b/(1+z_{abs})$ . The nine  $\lambda_{rb}$  values (column 4, Table 1b) have a mean of 918.5 Å ( $\sigma = 2.7$  Å) and show an intrinsic scatter which is greater than the measurement errors.

The value of  $\lambda_{rb}$  is expected to increase with: (1) H I gas velocity dispersion; (2) column density  $N$  (H I) and (3) instrumental resolution  $\Delta\lambda$ ; all in a predictable manner. Knowing  $N$  (H I) and  $\Delta\lambda$  we have a new simple and accurate method of determining velocity dispersions in optically thick hydrogen (D.T. and R. F. Carswell, in preparation).

Recalling that systems observed through the presence of absorption lines will be biased against low  $\lambda_b$  and  $\lambda_{rb}$  values,  $\lambda_{rb} = 918.0$  Å is chosen as a mean value likely to apply to all LLS. The expected instrumental resolution dependence is negligible for the present sample. The redshifts for the LLS are defined as:  $z_{LLS} = (\lambda_b/918) - 1$ ; or  $z_{LLS} = z_{abs}$  if  $z_{abs}$  is known and listed in Table 1b. Similarly we have  $z_{min} = (\lambda_{min}/918) - 1$  as the minimum redshift at which an LLS could have been observed.

## Distribution in velocity

In the intervening hypothesis, a non-evolving population of absorbers in a Friedmann cosmological model with zero cosmological constant  $\Lambda$  will give rise to an observed distribution of

$$N(z) = N_0(1+z)(1+2q_0z)^{-1/2}, \quad \Lambda = 0 \quad (1)$$

absorbers per unit redshift along a line of sight<sup>3,15</sup>, where  $N_0$  is  $N(z=0)$ . The velocity  $v$  of an absorption system of redshift  $z_{abs}$ , as measured from the rest frame of the QSO in whose line of sight the absorption occurs, is given by

$$\beta = v/c = (R^2 - 1)/(R^2 + 1), \quad (2)$$

$$R = (1+z_{em})/(1+z_{abs})$$

where the  $c$  is the speed of light and  $z_{em}$  is the QSO emission redshift. A positive  $\beta$  implies that the absorber is receding from the QSO. Combining equations (1) and (2) we can derive  $Y(\beta, z_{em})$ —the predicted number of absorbers per unit  $\beta$  along the line of sight to a QSO at  $z_{em}$ . For the intervening hypothesis, without evolution, it is found that

$$Y(\beta, z_{em}) = N(z_{abs}) \left| \frac{\partial z_{abs}}{\partial \beta} \right|_{z_{em}}^{-1} \\ = N_0(1+z_{em})^2(1+\beta)^{-2} \\ \times \left[ 1 + 2q_0 \left\{ \left[ \frac{1-\beta}{1+\beta} \right]^{1/2} (1+z_{em}) - 1 \right\} \right]^{-1/2} \quad (3)$$

Note the strong dependence on  $z_{em}$ . If the absorbers are intrinsic there is no *a priori* reason to expect  $Y(\beta)$  to depend on  $z_{em}$ , though in fact there is no clear prediction even for the  $\beta$  distribution of intrinsic absorption. I now outline statistical methods that can be used to estimate  $Y(\beta)$  for the LLS.

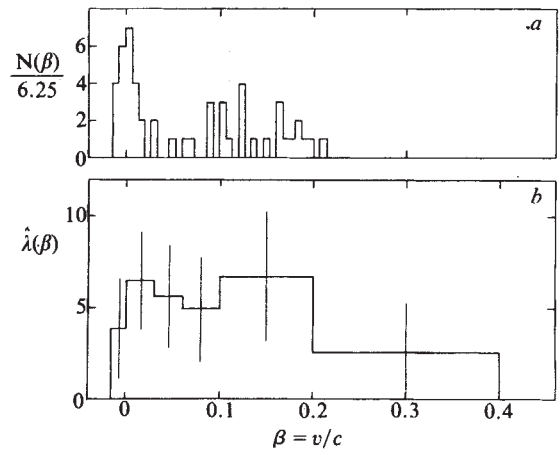
## Survival data analysis

The LLS data are in a form which is directly analogous to survival or failure time data encountered in biomathematics. Consider a general coordinate  $t$  which is some measure of the separation or distance from the QSO, to its LLS (for example,  $t = \beta$  or  $t = z_{em} - z_{abs}$ ). The maximum redshift for absorption systems is known to be close to  $z_{em}$ , which we equate with  $t(z_{abs} = z_{em}) = \beta = 0$ . For the  $i$ th QSO, an LLS could be observed at any value of  $t$  in the range  $t(0, T_i)$  where  $T_i = t(z_{min,i})$  is the censoring position corresponding to the minimum observed wavelength. I ignore for the present the occasional absorption systems with  $z_{abs} > z_{em}$  or  $\beta < 0$ , which have  $t < 0$ .

Three equivalent functions may be used to describe the distribution of LLS<sup>16</sup>:

(1) The observed density function,

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t < x < t + \Delta t)}{\Delta t} \quad (4)$$



**Fig. 2** Absorption system density as a function of velocity  $v = c\beta$  in the QSO rest frame. *a*, C IV systems from ref. 2 shown in histogram form.  $N(\beta)$ , the number of systems per unit  $\beta$  per QSO, has a mean of  $6.3 \pm 1.2$  for  $0.02 < \beta < 0.2$ . The total sample coverage is  $-0.02 < \beta < 0.22$ . *b*, The density  $\hat{\lambda}(\beta)$  of LLS per unit  $\beta$  per QSO, obtained using equation (9). The mean density for  $0 < \beta < 0.2$  is  $\hat{\lambda}(\beta) = 6.0 \pm 1.4$ . Unlike  $N(\beta)$ ,  $\hat{\lambda}(\beta)$  is obtained by considering a maximum of one system (the system with lowest  $\beta$ ) per QSO. Note that the data are inhomogeneous because we neglect the  $z_{em}$  dependence in equation (3).

is the probability that an LLS is observed in a given line of sight at  $t = x$ .

(2) The survival function,

$$S(t) = 1 - \int_0^t f(t') dt' \quad (5)$$

is the probability that no LLS occurs in the interval  $(0, t)$ .

(3) The population density function,

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t < x < t + \Delta t | x > t)}{\Delta t} \quad (6)$$

is the probability that an LLS is observed in the interval  $(t, t + \Delta t)$ , conditional on the absence of LLS in the interval  $(0, t)$ . Because we have specified that a maximum of one LLS may be observed from each line of sight,  $f(t)$  is the observed density (per unit  $t$ ), while  $\lambda(t)$  is the density of the population from which the sample was drawn. If the intervening hypothesis is correct  $\lambda(\beta) = Y(\beta)$  for fixed  $z_{em}$ ; provided that the observed LLS are a fair sample of all LLS.

The simplest distribution for the parent population of LLS is a random distribution with constant density  $\lambda(t) = \lambda$ , independent of  $t$ . This is equivalent to a Poisson distribution for the number of LLS per unit  $t$ . The probability of no LLS in the interval  $(0, t)$  simplifies to<sup>16</sup>

$$S(t) = \exp \left[ - \int_0^t \lambda(t') dt' \right] = \exp(-\lambda t), \quad t > 0 \quad (7)$$

giving a (negative) exponential distribution for the observed LLS:

$$f(t) = \lambda \exp(-\lambda t), \quad t > 0 \quad (8)$$

The survival function may readily be estimated<sup>16,17</sup> using the Kaplan-Meier or product limit estimator  $\hat{S}(t)$ .  $\hat{S}(\beta)$  has been calculated for the sample of 20 LLS from the 36 QSOs given in Table 1a. Figure 1a shows the 'cumulative hazard'  $-\log_e \hat{S}(\beta)$  as a function of  $\beta$ , with asymptotic  $1\sigma$  error bars derived following Kalbfleisch and Prentice<sup>17</sup>. It is evident that the straight line,  $-\ln \hat{S}(\beta) \approx \lambda\beta + 0.09$ ,  $\lambda = 5.5$  is a good representation of the data for  $-0.01 \leq \beta \leq 0.2$ .

The constancy of the gradient  $\lambda$  implies that  $f(\beta)$  is indeed of the exponential form of equation (8). The maximum likelihood method may then be used to estimate  $\lambda$ . Note that  $t_i$  is the position of the LLS in the  $i$ th QSO spectrum and  $T_i$  is the maximum possible value for  $t_i$ . Let the indicator  $\delta_i = 1$  if an LLS is seen and  $\delta_i = 0$  otherwise. Then the maximum likelihood estimator for the density parameter  $\lambda$  in equation (8) is

Table 2 The density of LLS as a function of  $z_{em}$ 

| QSO group  | $z_{em}$   | $\langle(1+z_{em})^2\rangle$ | $n$ | $m$ | $\hat{\lambda}(\beta)$ | $\sigma(\hat{\lambda}(\beta))$ | $\frac{\hat{\lambda}(\beta)}{\langle(1+z_{em})^2\rangle}$ | $\hat{\lambda}(\psi)$ | $\sigma(\hat{\lambda}(\psi))$ | $\langle M_v(q_0=0) \rangle$ | $\langle M_v(q_0=1) \rangle$ |
|------------|------------|------------------------------|-----|-----|------------------------|--------------------------------|-----------------------------------------------------------|-----------------------|-------------------------------|------------------------------|------------------------------|
| 1          | 0.533–2.20 | 6.52                         | 7   | 4   | 2.5                    | 1.2                            | $0.39 \pm 0.19$                                           | 0.48                  | 0.24                          | –27.7 (1.7)                  | –26.6 (1.4)                  |
| 2          | 2.69–2.90  | 14.63                        | 10  | 2   | 3.8                    | 2.6                            | $0.26 \pm 0.18$                                           | 0.27                  | 0.19                          | –28.3 (0.6)                  | –26.4 (0.7)                  |
| 3          | 3.00–3.53  | 17.42                        | 17  | 12  | 9.3                    | 2.7                            | $0.54 \pm 0.16$                                           | 0.59                  | 0.17                          | –28.2 (0.9)                  | –26.1 (0.9)                  |
| All QSOs   | 0.533–3.53 | 14.35                        | 34  | 18  | 5.3                    | 1.2                            | $0.37 \pm 0.08$                                           | 0.50                  | 0.12                          | –28.1 (1.1)                  | –26.3 (1.0)                  |
| $\chi^2_2$ |            |                              |     |     | 7.8                    |                                | 1.5                                                       | 1.7                   |                               |                              |                              |

Data from Table 1a. The two QSOs 0528–250 ( $z_{em}=2.765$ ) and 1402+044 ( $z_{em}=3.20$ ) which have  $z_{abs} > z_{em}$  are omitted because equation (9) is used to obtain  $\hat{\lambda}(t)$ , the estimate of the density of systems per unit  $t$ , per QSO.  $\langle x \rangle$  is the mean of  $x$ .  $n$  is the number of QSOs in a group. They have a total of  $m$  LLS.  $\psi$  is evaluated from equation (13) for  $q_0=0$ . Absolute magnitudes  $M_v$  for a Hubble constant of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  are from Table 1a. The two QSOs with  $z_{abs} > z_{em}$  are included but two other QSOs (0103.9–2913,  $z_{em}=2.80$  and 0244.4–3017,  $z_{em}=3.103$ ) are omitted, because we lack magnitude estimates. Figures in parentheses,  $M_v$ , ( $\sigma(M_v)$ ), are the sample standard deviations. About 25% of the magnitudes in Table 1a carry errors of 0.5–1 magnitudes.  $\chi^2$  values are for the three redshift groups, giving 2 d.f.  $\text{Prob}(\chi^2_2 > 1.7) = 43\%$  and  $\text{Prob}(\chi^2_2 > 7.8) = 2\%$ .

$$\hat{\lambda} = m / \sum [\delta_i t_i + (1 - \delta_i) T_i] \quad (9)$$

where a total of  $n$  QSOs were observed and  $m = \sum \delta_i$  is the total number of LLS in the sample. Bartholomew<sup>18</sup> has shown that for large  $n$ ,

$$\sigma(\hat{\lambda}) \approx \hat{\lambda} / \left[ \sum P_i \right]^{1/2}, \quad P_i = 1 - \exp(-\hat{\lambda} T_i) \quad (10)$$

where  $P_i$  is the probability that QSO<sub>*i*</sub> will show a LLS, on the condition  $\lambda = \hat{\lambda}$ . For large samples  $\sum P_i \approx m$ .

Application of equation (9) to the LLS data for  $t = \beta > 0$  only gives  $\hat{\lambda} = 5.3 \pm 1.2$  LLS per unit  $\beta$ , per QSO. (We omit the QSOs 0528–250 and 1402+044 which have  $z_{abs} > z_{em}$  or  $t < 0$  from all applications of 9.) *A posteriori* censoring of all data at  $\beta = 0.2$  gives  $\hat{\lambda} = 6.0 \pm 1.4$ . These two fits are shown on Fig. 1a. There is apparently some evidence that  $\lambda(\beta)$  decreases with increasing  $\beta$ , as would be expected for both the intervening and intrinsic hypotheses, but I do not present a formal test as the data are in fact inhomogeneous because QSOs of different  $z_{em}$  have different  $\lambda_{min}$  or  $T_i$  values (see equation (3)).

Figure 2b shows explicitly  $\hat{\lambda}(\beta)$  as a function of  $\beta$ . The density of systems per unit velocity is approximately constant for  $-0.01 \leq \beta \leq 0.2$ . There is certainly no massive congregation of LLS at  $\beta = 0$ , such as that claimed by Weymann *et al.*<sup>2</sup> for a mixed sample of narrow metal line systems which have conspicuous C IV absorption lines (Fig. 2a).

Recent homogeneous surveys of Ly $\alpha$ , C IV and Mg II absorption systems<sup>3,5,6</sup> have shown that in no case is there a massive congregation of systems at  $\beta = 0$ , evidence which supports the intervening hypothesis, as the magnitude of the previously claimed excess was an embarrassment to that hypothesis<sup>11</sup>. The apparent excess at  $\beta \approx 0$  had the same velocity width as QSO emission lines and was probably a consequence of the decrease in the minimum detectable equivalent width of absorption lines which lie in emission lines<sup>5,6</sup>. In addition I strongly suspect that the claimed excess of systems<sup>19,20</sup> with  $R = (1 + z_{em})/(1 + z_{abs}) = 1.11$  (the so-called line or edge locking controversy<sup>2</sup>) is a further manifestation of the same data inhomogeneity. Systems with  $R = 1.107$  will have C IV  $\lambda\lambda$  1548, 1,550 absorption lines which lie in the centre of the Si IV + O IV  $\lambda$  1,400 QSO emission line, and so will be easier to detect.

It was claimed above that if the non-evolving intervening absorber hypothesis is correct  $\lambda(\beta)$  should follow  $Y(\beta, z_{em})$ , equation (3), and vary as  $(1 + z_{em})^2$  for  $q_0 = 0$ . To test this prediction we have divided the QSOs into three groups on the basis of  $z_{em}$  (Table 2).

Figure 1b shows that  $\lambda(\beta)$ , the mean density of LLS per unit  $\beta$ , is indeed far greater for QSOs with  $z_{em} \geq 3.0$  (solid lines) than for those QSOs with  $z_{em} \leq 2.2$  (dashed lines). Two formal tests of the significance of the  $z_{em}$  dependence of the LLS distribution have been applied. The logrank statistic<sup>21</sup> applied to the QSOs in the lowest and highest redshift groups gives a standard deviate of 2.07, an indication that the  $z_{em}$  dependence is significant at the 96% level. Gehan's statistic<sup>21</sup> coincidentally gives the same significance level.

These results are strong support for the intervening absorber hypothesis.

## Cosmological model

The redshift dependence of the LLS distribution will now be assessed using the coordinate  $z_{LLS}$  instead of  $\beta$ . Using the maximum likelihood method to estimate  $q_0$  in the distribution given by equation (1), it is found that the likelihood is maximized for  $q_0 = 0$ , with a  $1\sigma$  upper bound of  $q_0 = 3.1$ . Because  $q_0 = 0$  is the extremity of the range for  $q_0$  in equation (1), we try a more general model for the possible redshift dependence for the LLS distribution<sup>15</sup>;

$$N(z) \propto \lambda(z) = \lambda_0(1+z)^G \quad (11)$$

it is found that  $G = 1.2$  with  $1\sigma$  confidence bounds of  $G = 0.3$  and 2.3. From equation (1) we expect  $G = 1$  for  $q_0 = 0$ , and  $G = \frac{1}{2}$  for  $q_0 = \frac{1}{2}$ . The LLS data appear to be well represented by a non-evolving population of absorbers in a Friedmann model with zero cosmological constant and  $q_0 \sim 0$ .

This proposition can be tested explicitly by transforming to a coordinate  $X(z)$  in which the density of absorbers per unit  $X$  should be independent of  $X$  and  $z$  (ref. 22):

$$\lambda(X) = \lambda(z) |dz/dX| = \text{constant}$$

By setting  $X(z=0) = 0$  and using equation (1) then:

$$X(z) = \int_0^z (1+z')^{-1/2} dz', \quad \Lambda = 0$$

$$= \begin{cases} \frac{1}{2}[(1+z)^2 - 1], & q_0 = 0 \\ \frac{2}{3}[(1+z)^{3/2} - 1], & q_0 = \frac{1}{2} \end{cases} \quad (12)$$

It is convenient to examine the distribution of the LLS using the coordinate  $\psi$ :

$$\psi(z) = X(z_{em}) - X(z) \quad (13)$$

to represent  $t$  in equations (7)–(10). Figure 3a shows the survival function for the LLS in the coordinate  $\psi$  evaluated for  $q_0 = 0$ . The data are well represented by the exponential model (equations (7) and (8)) with a density  $\hat{\lambda}(\psi) = 0.50$ ,  $\sigma(\hat{\lambda}) = 0.12$ .

Figure 3b shows the survival function for the 18 QSOs with  $z_{em} \geq 3$  (solid lines) and 7 QSOs with  $z_{em} \leq 2.2$  (dashed lines). A comparison with Fig. 1b, which illustrates the same data, using  $\beta$  as the coordinate, shows that  $\psi$  has successfully accounted for the  $z_{em}$  dependence in the LLS distribution. The slight remaining difference between  $\hat{S}(\psi)$  for QSOs of different  $z_{em}$  is not statistically significant, as the  $\hat{\lambda}(\psi)$  values in Table 2 indicate.

## Sample homogeneity

Does the exponential model with  $\lambda(\psi(q_0=0)) = 0.5$  fully account for the LLS distribution? One possibility is that the individual QSOs might be ionizing the LLS in their lines of sight, producing a deficit of LLS close to the QSOs. The distribution of LLS would then depend on the QSO absolute magnitudes, which are given in Table 1a.

Figure 3c shows the survival function for the LLS, with the QSOs in two groups, depending on whether their  $q_0 = 0$  absolute magnitudes are less than or greater than the mean absolute magnitude of  $M_v = -28.13$ . The gradient  $\lambda(\psi)$  of cumulative



hazard  $-\ln \hat{S}(\psi)$  is larger for the less luminous QSOs (dashed lines). The difference is not statistically significant: the logrank test yields a standardized normal statistic of 1.12. A preliminary analysis of an augmented sample containing 70 QSOs and 30 LLS argues against any absolute magnitude dependence (D.T. and M. A. J. Snijders, in preparation).

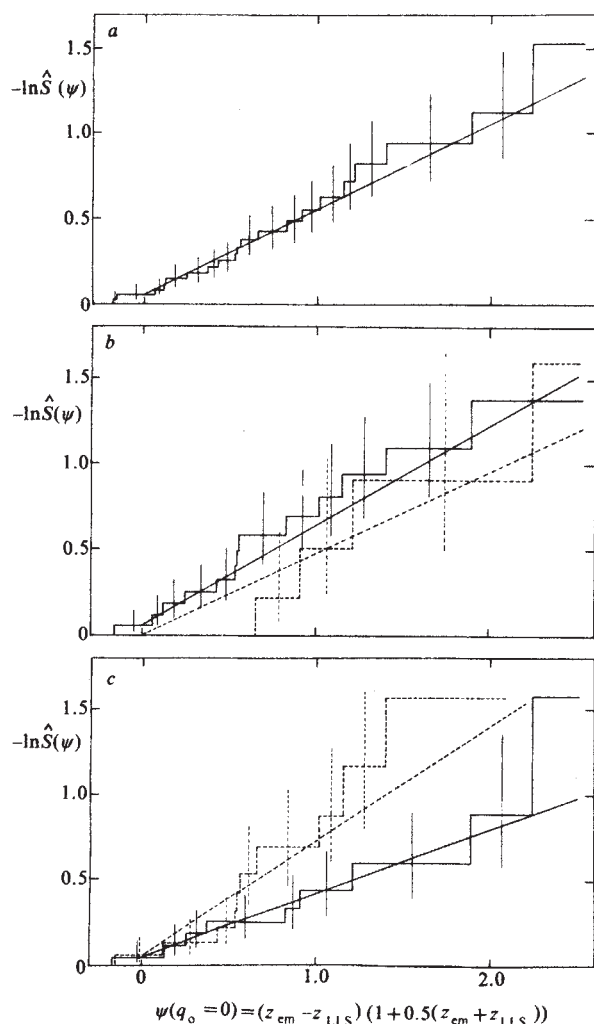
Given the set of censoring positions,

$$T(z_{\min,1}), \dots, T(z_{\min,n})$$

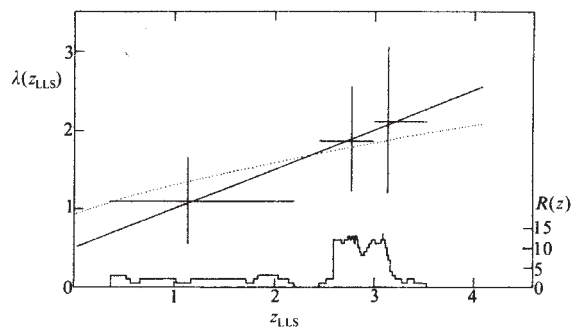
for the  $n$  QSOs in the sample we can apply statistical 'goodness of fit' tests which are sensitive to the variance of the positions of the  $m$  observed LLS:

$$t_1, \dots, t_m \text{ and } \delta_1, \dots, \delta_n, \quad m = \sum \delta_i \leq n$$

where  $t_i = t(z_{\text{LLS},i})$  and  $\delta_i = 1$  if an LLS is observed in the spectrum of the  $i$ th QSO, or  $\delta_i = 0$  otherwise. We are specifically interested in assessing two possibilities. First, that QSO sample selection bias will have occurred. Observers may have chosen to study QSOs which, on the basis of prior information (absorption spectra which do not cover LLS wavelengths), are likely or unlikely to have an LLS. Second, the data may be intrinsically inhomogeneous. Although the overall sample of LLS may apparently be well represented by a single exponential distribu-



**Fig. 3** Cumulative Hazard plots for the LLS sample as a function of  $\psi(q_0=0)$ , equation (13). *a*, The straight line is  $-\ln \hat{S}(\psi) = \lambda \psi$  with  $\hat{\lambda} = 0.50$ ,  $\sigma(\hat{\lambda}) = 0.12$ . *b*, QSO redshift division. The solid lines are for the 18 QSOs with  $z_{\text{em}} \geq 3.0$ . The straight solid line has the gradient  $\hat{\lambda} = 0.59$ ,  $\sigma(\hat{\lambda}) = 0.17$ . The dashed lines are for the seven QSOs with  $z_{\text{em}} \leq 2.2$  and are fitted by the density  $\hat{\lambda} = 0.48$ ,  $\sigma(\hat{\lambda}) = 0.24$ . *c*, QSO absolute magnitude division. Solid lines are for the 18 most luminous QSOs all with  $M_v < -28.13$  ( $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $q_0 = 0$ ). The straight line fit has the density  $\hat{\lambda} = 0.37$ ,  $\sigma(\hat{\lambda}) = 0.13$ . Dashed lines are for the 16 least luminous QSOs with  $M_v > -28.13$ . The fit has gradient  $\hat{\lambda} = 0.66$ ,  $\sigma(\hat{\lambda}) = 0.22$ .



**Fig. 4** LLS density as a function of redshift. The sample of 34 QSOs showing 18 LLS all with  $z_{\text{LLS}} < z_{\text{em}}$  are included. The two QSOs with  $z_{\text{LLS}} > z_{\text{em}}$  are excluded. Crosses indicate  $\hat{\lambda}(\langle z_{\text{LLS}} \rangle)$ , the density of LLS per unit  $z$ , per QSO with  $1\sigma$  asymptotic error bounds given by equation (10).  $\langle z_{\text{LLS}} \rangle$  is the *a priori* mean redshift sampled by the QSOs in each bin. The straight solid line is  $\lambda(z) = \lambda_0(1+z)$  with  $\lambda_0 = 0.50$ , the expected distribution of non-evolving absorbers in a Friedmann model with zero cosmological constant and  $q_0 = 0$ . The dotted curve is  $\lambda(z) = \lambda_0(1+z)^{1/2}$  with  $\lambda_0 = 0.93$ , appropriate for  $q_0 = \frac{1}{2}$ . In both instances  $\lambda_0 = \lambda(z=0)$  was estimated using the maximum likelihood method applied to the unbinned data. The lower histogram shows  $R(z)$ , the *a posteriori* risk set: the number of QSOs towards which a LLS could have been observed as a function of redshift. A QSO enters the risk set at  $z_{\text{em}}$  and leaves at  $z_{\text{LLS}}$ , or  $z_{\text{min}}$  if no LLS was observed.

tion; it might be the case that  $\lambda$  is a function of some concomitant variable. (Note that the  $z_{\text{em}}$  dependence of  $\lambda(\beta)$  was not apparent in Fig. 1*a*.)

A statistic which will be sensitive to the overall extent of both possibilities is:

$$Q_n = \sum_i \frac{(\ln L_i - E(\ln L_i))^2}{V(\ln L_i)} \quad (14)$$

$$L_i = \delta_i f(t_i; \theta) + (1 - \delta_i) S(T_i; \theta)$$

where  $L_i$  is the likelihood of the  $i$ th of the  $n$  QSO observations, and  $\theta = 1/\lambda$ . Because we are not specifically interested in the value for the scale parameter  $\theta$ , we transform the units of measurement of the observed values of the  $t_i$  and  $T_i$  to give  $\hat{\theta} = 1$  that is  $t'_i = t_i/\hat{\theta}$ ,  $T'_i = T_i/\hat{\theta}$ . The expectation  $E$  and variance  $V$  are evaluated over all possible  $t'_i(0, T'_i)$  and  $\delta_i = 0, 1$  for fixed  $T'_i$  and  $\theta' = 1$ .  $Q_n$  has an approximately  $\chi^2_{n-1}$  distribution, and should be interpreted in the same way as  $\chi^2$ . In terms of the  $\psi(q_0=0)$  coordinate we obtain  $Q_{34} = 33.5$  for the LLS data, while for  $\beta$ ,  $Q_{34} = 46.9$ . A Monte-Carlo experiment gave  $\text{Prob}[Q_{34}\{T(\psi)\} > 33.5] = 45\%$  and  $\text{Prob}[Q_{34}\{T(\beta)\} > 46.9] = 6\%$ . The  $Q$  test is clearly powerful enough to indicate the  $z_{\text{em}}$  dependence of  $\lambda(\beta)$ . We may then conclude that the sample selection bias is probably not serious and that the data are well represented by the exponential distribution with the single parameter  $\lambda(\psi)$ .

## Implications

I have argued that the redshift dependence of the Lyman limit absorption system (LLS) density is in agreement with that expected for a non-evolving population of intervening absorbers. It seems highly unlikely that the density of intrinsic systems would depend on the QSO redshifts and yet be constant for velocities (measured in the QSO rest frames) in the range  $-0.01 \leq \beta = v/c \leq 0.2$ .

It would be easy to devise an *a posteriori* evolutionary scheme in which the mean velocities of intrinsic clouds increase with cosmic time or decreasing  $z_{\text{em}}$ . However, such models are unlikely to predict  $\lambda(z) \propto (1+z)$ , as was found for the LLS data, and would lead one to expect that low  $z_{\text{em}}$  QSOs should have absorption systems with large blueshifts.

Figure 4 shows explicitly the redshift dependence of the LLS density, although it must be stressed that much information is lost in this plot because the data are binned and  $\lambda(z)$  varies across the bins.

The density of absorption systems as a function of redshift contains valuable information on both the evolutionary behaviour of the absorbing material and the cosmological

model<sup>15</sup>. If we accept a model with cosmological constant  $\Lambda = 0$  and  $0 < q_0 \leq \frac{1}{2}$  as a good first choice, then Fig. 4 shows that the absorbers causing the LLS do not evolve significantly over the redshift range  $0.4 \leq z \leq 3.5$ , or  $\sim 50\%$  of the history of the Universe.

The LLS absorption may arise from material in the outer regions of galaxies or in intergalactic material. In either case there has apparently been little evolution since  $z \approx 3.5$ . This is a clear indication that galaxy formation occurred at  $z > 3.5$ , because gaseous protogalaxies with radii  $\sim 100$  kpc would be highly conspicuous in absorption; even though absorption by dust may make them inconspicuous at optical wavelengths. Alternatively, if essentially all the dust and gas has been processed (necessarily at  $z > 3.5$ ) into stars before collapse, then the protogalaxies are likely to be conspicuous in emission and should have been seen in recent searches<sup>23</sup>. Low redshift ( $z \leq 3$ ) protogalaxies will only be hard to detect if they are optically thin to absorption before collapse and become dusty and optically thick during collapse, concealing any bright phase of early star formation. In this case we should be looking for IR emission from sources of small angular size.

Received 26 January; accepted 11 May 1982.

1. Burbidge, G. *IAU Symp.* No. 92, 99–105 (1980); *Nature* **282**, 451–455 (1979).
2. Weymann, R. J., Carswell, R. F. & Smith, M. G. *Astr. Astrophys.* **19**, 41–76 (1981).
3. Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl.* **42**, 41–81 (1980).
4. Weymann, R. J., Williams, R. E., Peterson, B. M. & Turnshek, D. A. *Astrophys. J.* **234**, 33–46 (1979).
5. Young, P., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J.* (in the press).
6. Tytler, D., Boksenberg, A., Sargent, W. L. W. & Young, P. *Astrophys. J.* (submitted).
7. Bahcall, J. N. & Spitzer, L. *Astrophys. J. Lett.* **156**, L63–L65 (1969).
8. Bregman, J. N. *Astrophys. J.* **250**, 7–17 (1981).
9. Young, P., Sargent, W. L. W., Boksenberg, A. & Oke, J. B. *Astrophys. J.* **249**, 415–421 (1981).
10. Goldreich, P. *Phys. Scr.* **17**, 225–228 (1978).
11. Weymann, R. J. *IAU Symp.* No. 92, 107–117 (1980).
12. Smith, M. G. *et al. Mon. Not. R. astr. Soc.* **195**, 437–449 (1981).
13. Coleman, G. *et al. Astrophys. J.* **207**, 1–15 (1976).
14. Morton, D. C., Chen, J.-S., Wright, A. E., Peterson, B. A. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **193**, 399–413 (1980).
15. Tytler, D. *Nature* **291**, 289–292 (1981).
16. Gross, A. J. & Clark, V. A. *Survival Distributions* (Wiley, New York, 1975).
17. Kalbfleisch, J. D. & Prentice, R. L. *The Statistical Analysis of Failure Time Data* (Wiley, New York, 1980).
18. Bartholomew, D. J. *J. Am. stat. Ass.* **52**, 350–355 (1957).
19. Burbidge, E. M. & Burbidge, G. R. *Astrophys. J.* **202**, 287–295 (1975).
20. Drew, J. E. *Astr. Astrophys.* **66**, 343–354 (1978).
21. Prentice, R. L. & Marek, P. *Biometrics* **35**, 861–867 (1979).
22. Bahcall, J. N. & Peebles, P. J. E. *Astrophys. J. Lett.* **156**, L7–L10 (1969).
23. Ellis, R. S. *Lectures at the VIII course of the International School of Cosmology and Gravitation*, Erice (1981).
24. Davis, M. *IAU Symp.* No. 92, 57–64 (1980).
25. Sargent, W. L. W. *et al. Astrophys. J.* **230**, 49–67 (1979).

The present analysis shows that the QSOs and LLS are apparently separate, uncoupled populations. There is then every reason to expect the LLS to exist at  $z > 3.5$ , and perhaps continuing to show little evolution. This being the case I predict that little radiation (QSO, protogalaxy or otherwise) will be observed in the optical *U*, *B*, *V*, or *R* bands originating from  $z \geq 3.5$ , 4.2, 5.3 or 6.9, due simply to absorption by high redshift LLS. Such are the limits of optical astronomy.

This article has only considered the gross redshift distribution of the LLS. Their physical nature will be discussed elsewhere. I thank Alec Boksenberg, Bob Carswell, R. Galbraith, V. Isham, Keith Ponting and Toon Snijders for helpful discussions, and Alan Wright and Ray Wolstencroft for communicating results before publication. Financial support from the SERC is appreciated.

26. Osmer, P. S. *Astrophys. J.* **227**, 18–29 (1979).
27. Wolstencroft, R. D., Arp, H. C., Ku, W. H.-M. & Scarrott, S. M. Preprint (Royal Observatory of Edinburgh).
28. Green, R. F. *et al. Astrophys. J.* **239**, 483–494 (1980).
29. Grandi, S. A. & Tift, W. G. *Publ. astr. Soc. Pacif.* **86**, 873–884 (1974).
30. Wright, A. E., Peterson, B. A., Jauncey, D. L. & Condon, J. J. *Astrophys. J. Lett.* **226**, L61–L64 (1978).
31. Carswell, R. F., Strittmatter, P. A., Williams, R. E., Beaver, E. A. & Harms, R. *Astrophys. J.* **195**, 269–277 (1975).
32. Oke, J. B. *Astrophys. J. Lett.* **161**, L17–L20 (1970).
33. Chen, J.-S., Morton, D. C., Peterson, B. A., Wright, A. E. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **196**, 715–730 (1980).
34. Beaver, E. A. *et al. Astrophys. J. Lett.* **203**, L5–L8 (1976).
35. Snijders, M. A. J. *Proc. 2nd European Conf. on IUE*, Tübingen (ESA SP-157, 1980).
36. Boksenberg, A. *et al.* (in preparation).
37. Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A. & Tytler, D. *Mon. Not. R. astr. Soc.* **198**, 91–110 (1982).
38. Boksenberg, A. & Snijders, M. A. J. *Mon. Not. R. astr. Soc.* **194**, 353–364 (1981).
39. Young, P. J., Deverill, R. S., Gunn, J. E., Westphal, J. A. & Kristian, J. *Astrophys. J.* **244**, 723–735 (1981).
40. Snijders, M. A. J., Pettini, M. & Boksenberg, A. *Astrophys. J.* **245**, 386–395 (1981).
41. Peterson, B. A., Jauncey, D. L., Wright, A. E. & Condon, J. J. *Astrophys. J. Lett.* **222**, L81–L83 (1978).
42. Baldwin, J. A. *et al. Astrophys. J.* **193**, 513–523 (1974).
43. Jauncey, D. L., Wright, A. E., Peterson, B. A. & Condon, J. J. *Astrophys. J. Lett.* **223**, L1–L3 (1978).
44. Hewer, A., Boksenberg, A., Carswell, R. F., Smith, M. G. & Whelan, J. A. J. (in preparation).
45. Bregman, J. N., Glassgold, A. E. & Huggins, P. J. *Astrophys. J.* **249**, 13–16 (1981).
46. Boksenberg, A., Carswell, R. F. & Sargent, W. L. W. *Astrophys. J.* **227**, 307–379 (1979).
47. Hewitt, A. & Burbidge, G. *Astrophys. J. Suppl.* **43**, 57–158 (1980); **46**, 113–115 (1981).
48. Evans, A. & Hart, D. *Astr. Astrophys.* **58**, 241–245 (1977).
49. Wilson, R., Carnochan, D. J. & Gondhalekar, P. M. *Nature* **277**, 457–458 (1979).

# Intermontane-basin development in the past 4 Myr in the north-west Himalaya

Douglas W. Burbank\* & Gary D. Johnson

Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

*Through the combined use of magnetic-polarity stratigraphy with fission-track dating of volcanic ashes, a new chronology, spanning 4 Myr, has been developed for the intermontane basin of Kashmir in the northwestern Himalaya.*

SUPERIMPOSED intermontane basins in collisional tectonic settings typically develop during the late stages of an orogeny and are frequently transient phenomena. In the Himalaya, sedimentary basins, such as Peshawar, Campbellpore, Kashmir, Suketi and Kathmandu, evolve as tectonic depressions on the back side of lithic thrust-bounded sheets that constitute the schuppenstruktur on the outer, southern margin of the orogenic belt. As structural deformation migrates towards the bounding foredeep, continued uplift and erosion often obliterate the record of intermontane-basin sedimentation. Evaluating the sedimentary record within young basins in presently growing

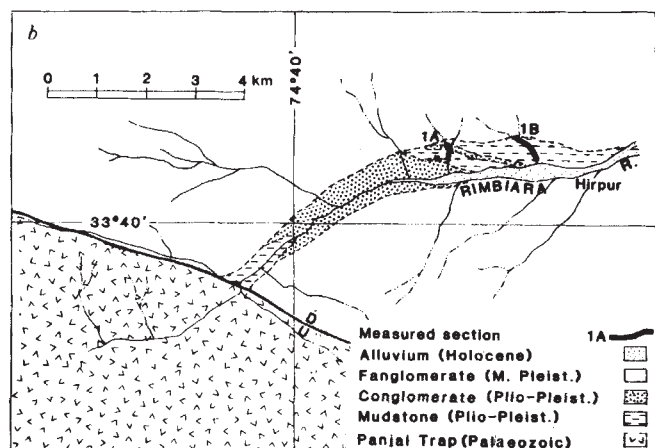
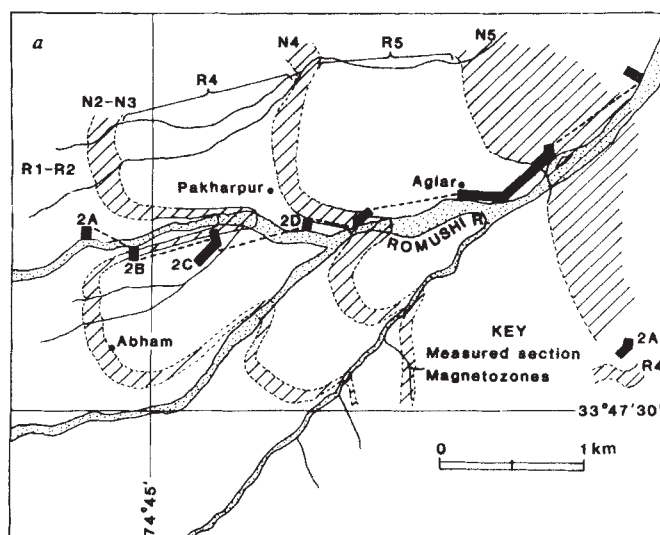
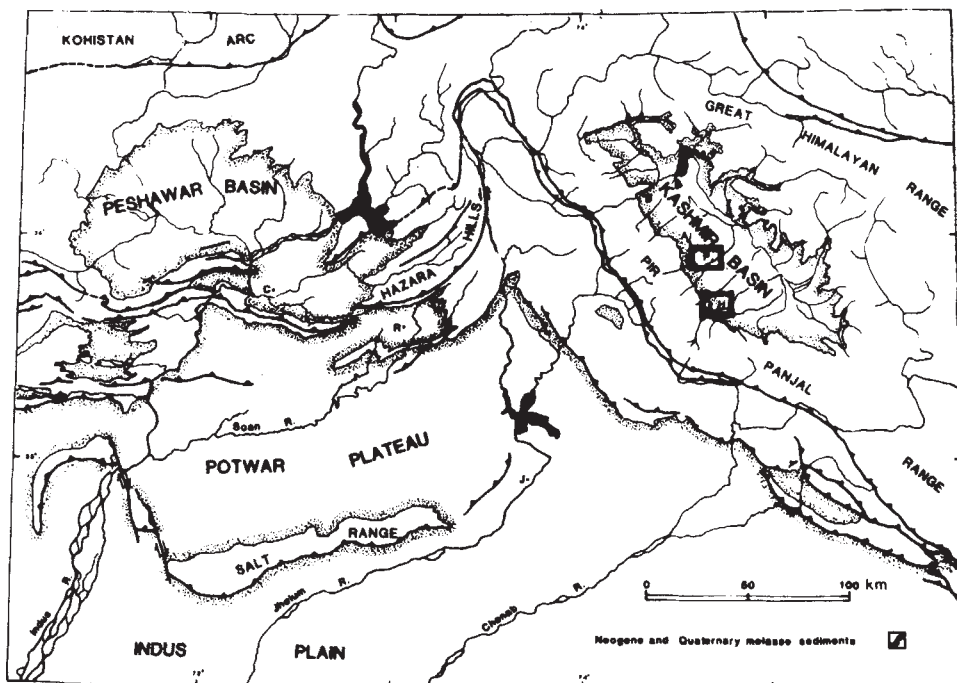
mountain ranges helps to delineate the history of basin evolution and place constraints on the tectonic activity of the enclosing structures.

The Kashmir Basin in the northwestern Himalaya (Fig. 1) is an intermontane basin presently experiencing rapid uplift along its southwestern margin. However, synorogenic sediments spanning nearly the entire period of basin development have not yet been eroded. These sediments, known collectively as the 'Karewas' or Karewa Group<sup>1–5</sup>, have been examined many times, because they contain late Cenozoic mammalian fossils<sup>2,6,7</sup> and a diverse invertebrate fauna<sup>8,9</sup>, they interfinger with glacial and periglacial deposits along the basin margin<sup>2,10–12</sup>, and they constitute a sedimentary sequence 1,300 m thick that potentially records the history of basin development<sup>1–5,12–17</sup>.

\* Present address: Department of Geological Sciences, University of Southern California, Los Angeles, California 90007, USA.



**Fig. 1** Map of the Kashmir and Peshawar intermontane basins of the northwestern Himalaya, symmetrically oriented around the North-west Syntaxis, a structural re-entrant defined by acute bends in the marginal thrusts (barbed lines). Movement along the imbricate thrusts on the southwestern margin of the Kashmir Basin has elevated the Pir Panjal Range and created the structural basin in which the intermontane sediments described herein have accumulated. A, Arigam; B, Baramula; C, Campbellpore; H, Hirpur; J, Jhelum; P, Pakharpur; R, Rawalpindi; S, Srinagar. Boxes show study areas.



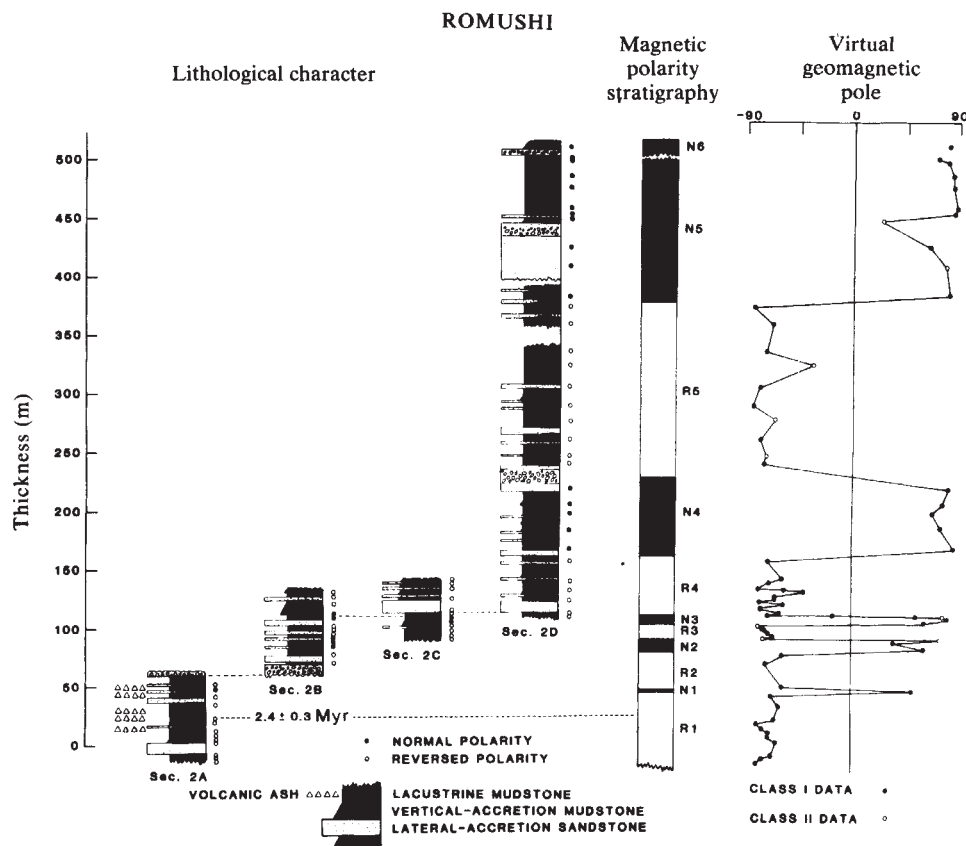
**Fig. 2** *a*, Location map of the Romushi section near Pakharpur, indicating measured sections and magnetozones (R1–N5). *b*, Map of exposures of the basal portions of the Karewa sequence along the upper Rimbiara River. The measured sections overlie >500 m of conglomerates and contorted mudstones, faulted against Palaeozoic bedrock.

Little agreement exists among earlier studies concerning the absolute chronology of the Karewa sequence. Recently, preliminary magnetic-polarity studies in portions of the Karewas along the southern end of the Kashmir Basin<sup>14</sup> have suggested that they span the late Gilbert to early Matuyama magnetic-polarity chrons, ranging from ~3.5 to 2 Myr ago. The Karewas have previously been interpreted as being solely Pleistocene in age<sup>2,7,11</sup>, as spanning the late Pliocene and much of the Pleistocene<sup>3,5,12,13,15,16</sup>, and as extending from perhaps late Miocene to the Pleistocene<sup>8,7</sup>. We describe here our efforts to date the Karewa sequence through the combination of detailed stratigraphical analysis, magnetic-polarity stratigraphy (MPS) and fission-track dating of enclosed volcanic ashes, and to correlate the Karewa sequence with the late orogenic, molasse sediments of the Himalayan foredeep to the south-west<sup>18–20</sup>.

### Magnetic stratigraphy and fission-track chronology

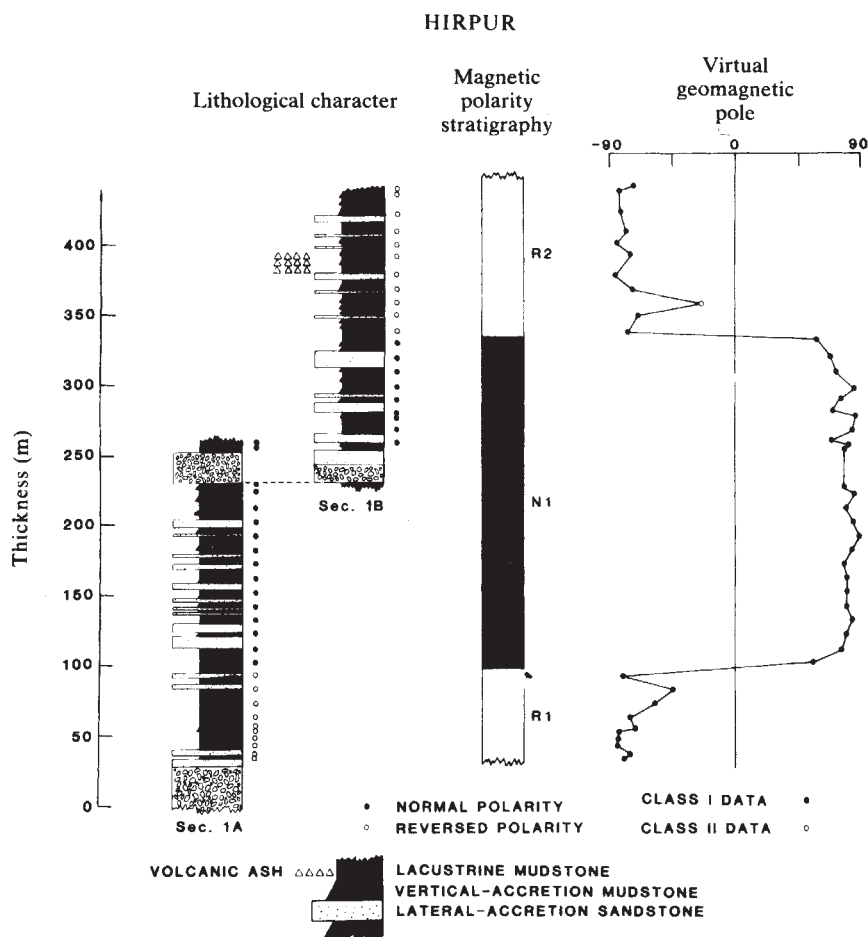
Rapid uplift of the Pir Panjal Range (Fig. 1), after deposition of the Karewas, has led to their dissection along the north-east flank of the range. More than 1,300 m of folded Karewa sediments are exposed along the Rimbiara and Romushi rivers in southwestern Kashmir (Figs 1, 2). Sixty-seven magnetic sites were established in these sediments along the Romushi River near Pakharpur (Fig. 2a), where an aggregate thickness of 530 m was measured. An additional 46 magnetic sites were established in a 425-m sequence along the upper Rimbiara River, upstream from Hirpur (Fig. 2b). Three, oriented samples were collected at each site and analysed following the techniques described in ref. 21.

Occasional, post-depositional magnetic overprinting observed in some samples was removed by demagnetization in an alternating field (a.f.) of 150–200 Oe. After such demagnetization, the magnetic vectors for 59 sites from the Romushi section were statistically significant ( $K > 10$ ) and yielded class I data<sup>20,22</sup>. The sample vectors of the eight other sites were more dispersed ( $K < 10$ ). For these latter sites (class II), the most deviant magnetic vector was disregarded when calculating the site virtual geomagnetic pole (VGP). The palaeolatitude of the site VGP forms the basis of the MPS of the Romushi section (Fig. 3). The dominantly reversely magnetized sediments (magnetozones R1–R5; Fig. 3) are punctuated by brief normal intervals (N1–N4) and capped by 140 m of normally magnetized rocks (N5 and N6).

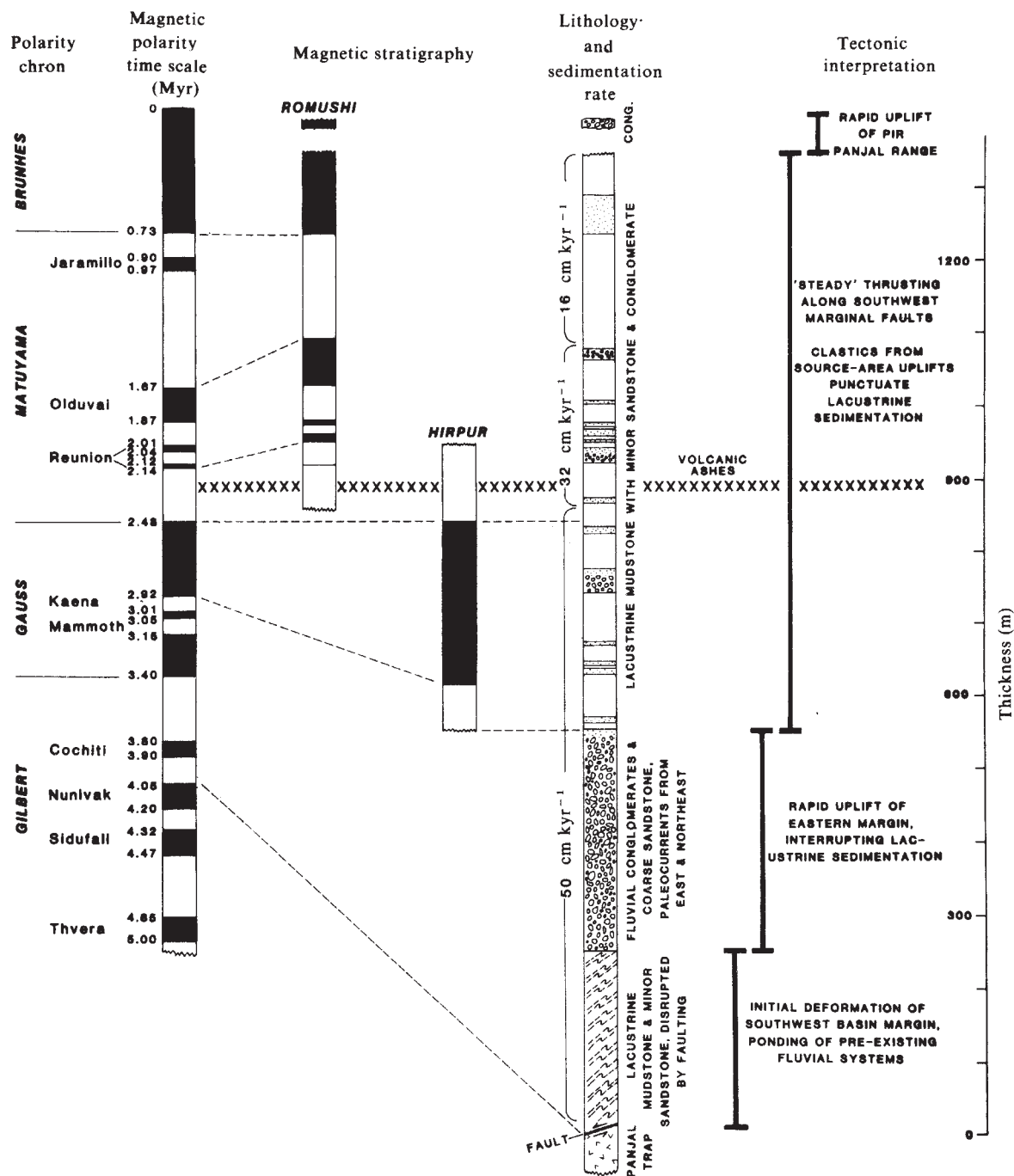


**Fig. 3** Lithological character and MPS of the Romushi section, Kashmir. MPS based on virtual geomagnetic pole plot of class I and II data. Volcanic ash at 35 m dated at 2.4 Myr. Magnetozone labels at right of MPS column. Interpretation: N2 and N3, Reunion subchrons; N4, Olduvai subchron; N5, Brunhes normal chron; R1-R5, Matuyama reversed chron.

**Fig. 4** Lithological character and MPS of the Hirpur section, Kashmir. Triplet of volcanic ashes at 380 m correlates with ash triplet at Romushi (30–40 m, section 2A, Fig. 3). Interpretation: R1, Kaena subchron; N1, upper Gauss normal chron; R2, lower Matuyama reversed chron.







**Fig. 5** Correlations of the magnetic-polarity time scale with the local MPS of Romushi and Hirpur place temporal controls on the record of intermontane-basin sedimentation. Depositional conditions are inferred to respond to tectonic deformation along the basin margin. Temporal constraints provided by the local MPS limit the duration of each tectonic interval. Extrapolation of sedimentation rates from the upper Gauss chron indicates that the basal mudstones at Hirpur were deposited at least 4 Myr ago. During the Matuyama chron, diminishing sedimentation rates precede the cessation of intermontane-basin sedimentation due to rapid uplift of the Pir Panjal Range.

Five volcanic ashes, some highly bentonitic, were discovered in the lowermost 70 m of the section (Fig. 3). Whereas the upper two ashes were thin (<0.5 cm) and discontinuous, the lower three, each 2–4 cm thick, span a 6-m interval and form an ash triplet. This triplet was recognized in the Dudhganga valley 10 km north-west of Pakharpur and the Rimbiara valley to the south-east. Zircons extracted from one of the lower ashes yielded a fission-track age of  $2.4 \pm 0.3$  ( $2\sigma$ ) Myr. Details of the magnetic and fission-track methodology and results are described elsewhere<sup>23</sup>.

This date indicates that magnetozones R1 and R2 represent the lower Matuyama reversed chron (Fig. 5). Magnetozones

N2, N3 and N4 are interpreted as depicting the Reunion and Olduvai subchrons. The upper Matuyama chron (R5) is followed by the Brunhes chron (N5 and N6). The Jaramillo subchron has not been found in the Romushi MPS, perhaps because of inaccessible, unsampled strata (340–360 m, section 2D, Fig. 3) or insufficient sample density. With this exception, the MPS from Romushi records all of the recognized polarity zones of the magnetic time scale<sup>24</sup> since 2.4 Myr ago. Comparison of the Romushi MPS with the magnetic time scale (Fig. 5) shows that sedimentation rates slowed after the Olduvai subchron. Consequently, N5 seems likely to represent at least half of the Brunhes magnetic chron.

The conglomeratic facies separating magnetozones N5 and N6 (Fig. 3) truncates folded Karewa sediments throughout south-west Kashmir. It records the rapid uplift of the Pir Panjal Range and concurrent cessation of widespread sedimentation in the Kashmir Basin during Brunhes time. Remnants of tilted, lacustrine Karewa strata above 3,150 m a.s.l. on the flanks of the Pir Panjal, 14 km west of Pakharpur, probably have been differentially uplifted a minimum of 1,500 m above their counterparts along the Romushi River since mid-Brunhes time. These data suggest from 4 to 10 mm yr<sup>-1</sup> of vertical uplift has occurred in the past 350,000 yr. Similar rates (~7 mm yr<sup>-1</sup>) for the nearby Nanga Parbat massif<sup>25</sup> suggest that broad areas of the northwestern Himalaya have undergone rapid uplift in the recent past. The post-Olduvai diminution of sedimentation rates at Romushi (Fig. 5) probably reflects earlier stages of uplift. In the Siwalik molasse, similar diminishing sedimentation rates have been caused by initial folding and progressive deformation of nearby foredeep structures before the termination of molasse sedimentation<sup>18</sup>.

Whereas the Romushi section spans most of the Brunhes and Matuyama magnetic chronos, the Hirpur section along the Rimbiara River (Fig. 2b) records an earlier period of intermontane-basin development. The sampled sections overlie more than 300 m of coarse conglomerates (Fig. 5), whose probable source lay to the east and north-east, as shown by palaeocurrent directions based on clast imbrications and crossbeds. These beds succeed an estimated 250 m of contorted mudstones and sandstones<sup>16</sup> that are faulted against Palaeozoic basement. Of the 46 magnetic sites (Fig. 4), 45 yielded class I data. Above the massive conglomerate near the base, the Hirpur MPS reveals a 70-m-thick sequence of reversed polarity (R1), followed by 238 m of normally magnetized rocks (N1). The upper 115 m of reversed rocks (R2) are truncated by the conglomerate analogous to that at the top of the Romushi section (N6; Fig. 3).

In the absence of other temporal constraints, the Hirpur MPS cannot be unambiguously placed within the magnetic-polarity time scale. However, a triplet of volcanic ashes, each 1–2 cm thick, was discovered 60 m below the top of the section (Fig. 4). These ashes are equivalent to the Romushi ash triplet discussed above, and were also deposited during the lower Matuyama chron (R2; Fig. 4). The two ash triplets bracket similar thicknesses of Karewa sediments and are nearly identical in appearance. At the three sites in Kashmir where the triplet was identified, there is a progressive thinning of the individual ashes from north-west to south-east, probably due to downwind attenuation. These Kashmir ashes have similar ages to many of the late Pliocene and Pleistocene ashes found in the Siwalik Group in northern Pakistan and India<sup>26</sup>. Their probable source lay to the west in the Dasht-e Nawar volcanic complex in central-eastern Afghanistan<sup>26</sup>.

Agrawal *et al.*<sup>14</sup> have recently suggested that the normally magnetized sediments in the vicinity of Hirpur (analogous to our N1) probably encompass the entire Gauss chron. However, we did not uncover reversed intervals representative of the Kaena and Mammoth subchrons in this rather densely sampled section (Fig. 4). The presence of only two magnetic reversals

among these 46 sites implies that the section spans a relatively brief time interval. In quantitative terms, when the random sampling model of McGee and Johnson<sup>27</sup> is applied to this palaeomagnetic data set, the time interval represented by this section should be <500,000 yr. Therefore, we interpret N1 as the uppermost part of the Gauss chron, not the entire Gauss. R1 would represent a portion of the Kaena subchron (Fig. 5).

Extrapolation of the sedimentation rate (50 cm kyr<sup>-1</sup>) at the sampled sequence at Hirpur to the underlying 550 m of conglomerates and mudstones yields an estimated age of 4 Myr for the base of the sequence. No older sediments attributable to the Karewa Group have been located in the Kashmir Basin. If the sedimentation pattern in Kashmir is similar to that observed in the Siwalik molasse<sup>18–20</sup>, the rate would be expected to accelerate in the early basin-filling stages and then decelerate in the waning phases as deformation intensified. Hence, the extrapolation of the maximum observed rate of 50 cm kyr<sup>-1</sup> to the base of the sequence at Hirpur probably underestimates the true age of the initiation of basin development.

## Conclusions

The Romushi and Hirpur sections, including the underlying strata at Hirpur, encompass virtually the entire interval of intermontane-basin sedimentation in Kashmir (Fig. 5). Initial ponding of pre-existing fluvial systems led to deposition of the basal Hirpur mudstones some 4 Myr ago. Strong uplift along the eastern basin margin caused an influx of conglomerates during the lower Gauss chron. This influx probably initiated in the upper Gilbert chron. Widespread lacustrine sedimentation began ~3 Myr ago and continued until ~0.3–0.4 Myr ago. During the lower Matuyama chron, volcanic ashes, probably erupted from the Dasht-e Nawar volcanic complex to the west, were deposited in coherent strata that thin to the south-east.

This new chronology from Kashmir demonstrates that, since ~4 Myr ago, intermontane-basin development, Karewa sedimentation and deposition of the Siwalik molasse to the south of the Pir Panjal Range have been concurrent. Deformation of the Siwalik molasse in the Potwar Plateau and the Jhelum re-entrant<sup>18–20</sup> during the Plio-Pleistocene is, in part, a response to continued thrusting and uplift of the south-west edge of the Kashmir basin along the marginal faults south of the Pir Panjal (Fig. 1). Whereas thrusting deforms the proximal foredeep margin south of the faults, it enhances lacustrine sedimentation in the intermontane basin by raising the local base-level. Consequently, the long lacustrine interval in the Karewa sequence between 3 and 1 Myr probably does not reflect quiescence, but is rather a response to continued tectonism along the basin margin. During the past 350,000 yr, very rapid uplift of the Pir Panjal Range has terminated widespread sedimentation in the Kashmir Basin.

This research was supported by grants from the Geological Society of America, Mobil Corp., and the NSF (INT 8019373 and EAR 8018779) to Dartmouth College. Logistical support from personnel of the Physical Research Laboratory (Ahmedabad), Delhi University, and the University of Peshawar is gratefully acknowledged.

Received 11 March; accepted 4 May 1982.

- Godwin-Austin, H. H. Q. *Jl geol. Soc. Lond.* **15**, 221–229 (1859).
- Terra, H. de & Paterson, T. *Carnegie Instn Washington Publ.* **493**, 80–251 (1939).
- Farooqi, I. A. & Desai, R. N. *J. geol. Soc. India* **15**, 299–305 (1974).
- Bhatt, D. K. *Himalayan Geol.* **6**, 197–208 (1976).
- Bhatt, D. K. & Chatterji, A. K. *Geol. Surv. India misc. Publ.* **41**, 191–200 (1979).
- Badam, G. L. *Curr. Sci.* **41**, 529–530 (1972).
- Tewari, B. S. & Kachroo, R. K. *Recent Res. Geol.* **3**, 468–477 (1977).
- Roy, D. K. *Geol. Surv. India misc. Publ.* **24**, 204–221 (1975).
- Singh, D. *Bull. Ind. geol. Ass.* **7**, 99–122 (1974).
- Kusumgar, S., Agrawal, D. P. & Krishnamurthy, R. V. *Radiocarbon* **22**, 757–762 (1980).
- Vishnu-Mittre *Geol. Surv. India misc. Publ.* **15**, 89–96 (1972).
- Wadia, D. N. *Int. geol. Congr., London* **10**, 43–48 (1951).
- Lydekker, R. *Rec. geol. Surv. Ind.* **11**, 30–69 (1876).
- Agrawal, D. P., Bhatt, D. K., Kusumgar, S. & Pant, P. K. *Proc. Ind. Acad. Sci.* **90**, 111–123 (1981).

- Bhatt, D. K. *Geol. Surv. Ind. misc. Publ.* **24**, 188–204 (1975).
- Bhatt, D. K. & Chatterji, A. K. *Proc. VI Ind. Colloq. Micropaleontological Stratigraphy*, 11–21 (1976).
- Singh, D. *Recent Res. Geol.* **3**, 79–86 (1977).
- Johnson, G. D., Johnson, N. M., Opdyke, N. D. & Tahirkheli, R. A. K. in *Geodynamics of Pakistan*, 149–165 (eds Farah, A. & deJong, K.) (Geological Survey of Pakistan, Quetta, 1979).
- Opdyke, N. D. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **27**, 1–34 (1979).
- Johnson, N. M. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37**, 17–42 (1982).
- Johnson, N. M., Opdyke, N. D. & Lindsay, F. H. *Bull. geol. Soc. Am.* **86**, 5–12 (1975).
- Fisher, R. A. *Proc. R. Soc. A* **217**, 295–305 (1953).
- Burbank, D. W. thesis, Dartmouth College (1982).
- Mankinen, F. A. & Dalrymple, G. B. *J. geophys. Res.* **84**, 615–627 (1979).
- Zeitler, P. K., Johnson, N. M., Tahirkheli, R. A. K. & Naeser, C. W. *Geol. Soc. Am. Abstr. Prog.* **13**, 587 (1981).
- Johnson, G. D. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37**, 63–93 (1982).
- McGee, V. F. & Johnson, N. M. *Trans. Am. geophys. Un.* **61**, 215 (1980).



# A glacio-isostatic facies model and amino acid stratigraphy for late Quaternary events in Spitsbergen and the Arctic

G. S. Boulton\*, C. T. Baldwin\*, J. D. Peacock†, A. M. McCabe‡, G. Miller§, J. Jarvis||, B. Horsefield\*, P. Worsley¶, N. Eyles\*, P. N. Chroston\*, T. E. Day\*, P. Gibbard#, P. E. Hare\*\* & V. von Brunn††

\* School of Environmental Sciences, University of East Anglia, Norwich NR4 7IJ, UK

† Institute of Geological Sciences, Murchison House, West Mains Road, Edinburgh EH9 3LA, UK

‡ Department of Geography, Ulster College, North Ireland Polytechnic, Newton Abbey, County Antrim BT37 0QB, UK

§ Institute of Arctic and Alpine Research, University of Colorado, Colorado 80302, USA

|| Department of Geography, University of St Andrews, St Andrews, Fife KY16 9AJ, UK

¶ Department of Geography, University of Reading, Whiteknights, Reading RG6 2AH, UK

# Department of Botany, Cambridge University, Downing Street, Cambridge CB2 3DY, UK

\*\* Department of Geophysics, Carnegie Institute, Washington, DC 20005, USA

†† Department of Geology, University of Natal, Pietermaritzburg, South Africa

*Strong early Holocene crustal uplift over the Spitsbergen archipelago reflects disappearance of an ice mass centred just north-east of the archipelago and which decayed rapidly just after 10,000 BP after a maximum extension between 12,600 and 10,000 BP. From the lithofacies distributions associated with these events a glacio-isostatic facies model can be defined which, combined with analyses of amino acid racemization, provides a valuable stratigraphic method for correlating pre-Holocene glacio-isostatic events. These are correlated with events in Greenland, and both are contrasted with Weichselian glacial events at the southern margins of the two great Northern Hemisphere ice sheets and with contemporary water mass changes in the North Atlantic and Norwegian Sea. The geographical variation in glacier response through time is explained by three principal climatic circulation patterns.*

DATING of raised shorelines<sup>1-7</sup> in the Spitsbergen archipelago allows the pattern of shoreline displacement during Holocene times to be reconstructed. It is usually supposed that isobase patterns such as that in Fig. 1a in areas known to have suffered earlier glacier expansion are primarily produced by isostatic uplift of crust previously depressed by glacier loading. On this assumption Fig. 1a implies a pre-Holocene ice mass centred east of the north-east sector of the archipelago.

At several sites (such as sites 26, 27 and 28 in Fig. 2) a phase of deep water mud deposition is shown to have commenced between 9,400 and 10,000 yr BP immediately after deposition of a till, thus locally dating deglaciation. During this time and for much of the rest of the Holocene, a dated sequence of raised beaches shows a record of falling sea level, which causes deep water mud sedimentation to be replaced by beach formation and eventually by sub-aerial exposure. These sedimentary sequences provide strong evidence for a glacial event which terminated 9,000 and 10,000 yr BP referred to as the Billefjord Stadial<sup>4</sup>. This caused isostatic depression of the crust; when the glacier retreated and the site was flooded, deep water muds accumulated; isostatic crustal uplift tends to lag behind unloading, but the accelerating uplift rate soon caused sea-level regression and beach formation and emergence. At any one site where a Holocene raised beach sequence has been dated it is possible to determine the depth of water throughout the Holocene at a point which is now at sea level. Thus at sites where there are dated marine deposits lying below the dated beach 'staircase', it is possible to assess the water depth in which the deposits formed. This has been done for several sites in a transect across the Holocene uplift (Fig. 1b). Although the end of the Billefjord Stadial can be clearly defined, dating its commencement is more difficult. At Kap Ekholm in Billefjord (Fig. 2, site 26) the stadial peak can be defined by tills which contain derived shells <sup>14</sup>C-dated at 11,028 yr BP, and which underlie marine beds dated at 9,709 yr BP. In several areas the till belonging to this stadial overlies sub-aerially weathered surfaces exposed since mid-Weichselian times.

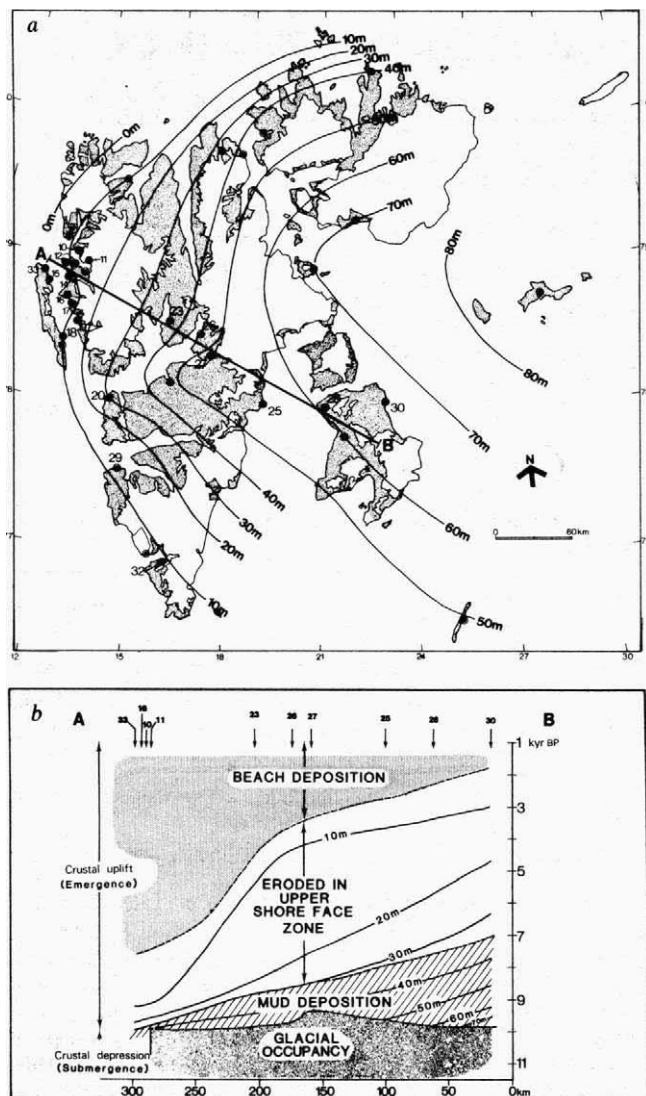
Large areas in western Spitsbergen were not overridden by Billefjord Stadial glaciers (Fig. 2, sites 33, 14 and 16)<sup>3,8</sup>, but

even there glacio-isostatic depression of the crust caused marine inundation. At Aavatsmarkbreen (Fig. 2, site 16) the Holocene regressive sequence which reflects crustal uplift is underlain by a fining upward sequence which we interpret as evidence of a marine transgression. This transgression was underway by 12,680 yr BP. We suggest that this reflects sea-level rise produced by crustal depression associated with expanding Billefjord Stadial glaciers, to be followed by regression caused by crustal uplift associated with glacier retreat. Thus, we date the Billefjord Stadial glacial maximum between 12,680 and 10,000 yr BP. The fact that here and elsewhere (Fig. 2) beds of this marine transgression lie above beach sediments of mid-Weichselian age rather than above a glacio-isostatic transgressive/regressive sequence related to glaciation at 18,000–20,000 yr BP leads us to conclude that there was no significant ice expansion at 18,000–20,000 yr BP the period of greatest Weichselian ice sheet expansion in Europe and North America.

This work allows us to demonstrate a very important temporal and spatial pattern of 'glacio-isostatic' facies development associated with glacier advance and retreat, which we believe to have general application (Fig. 1b). Glacier advance produces crustal depression. Submergence occurs beyond the glacier margins, and on glacier retreat the glaciated area is also inundated until accelerating uplift causes regression in both the previously glaciated area and just beyond the former ice margin. This uplift produces rapid displacement of facies so that glacio-marine mud sequences reflect short periods of time and suffer strong erosion in the sub-littoral zone (Fig. 1b). These mud sequences are thicker, reflect longer periods of time and reach higher elevations nearer to the centre of uplift.

## Glacio-isostatic facies model of pre-Holocene events

At many isolated localities, pre-Billefjord Stadial sediments occur which reflect several glacial and marine events. It is difficult to correlate and date such sequences in a continental area, where geological events tend to be represented by temporally and spatially discontinuous sediment lenses. However,



**Fig. 1** a, Elevation of the 9,500 yr BP isobase over the Spitsbergen archipelago. ○, Data points; ●, sites in Figs 1b and 3b. Line A-B is the transect shown in b. b, The changing depth of water from the Billefjord Stadial to the present day at points now at sea level in a transect across the archipelago. The character of sedimentation through time along transect A-B is shown.

the pre-Billefjord Stadial sediments tend to occur in repetitive till-mud-sand facies<sup>4</sup> sequences similar in character and thickness to the latest Weichselian-Holocene glacio-isostatic facies sequences shown in Fig. 1b. If we apply this model to the pre-Holocene sequences, we expect regionally extensive transgressive/regressive couplets which reflect glacio-isostatic events. The transgressive component is generally represented by a glacial sediment and the regressive component by dateable marine sediments which reflect short time periods and which can be correlated over large distances.

We couple this model with studies of the degree of amino acid racemization in fossil molluscs (*Mya truncata* and *Hiatella arctica*) commonly found *in situ* in the marine beds, to produce what we suggest to be a powerful stratigraphic method.

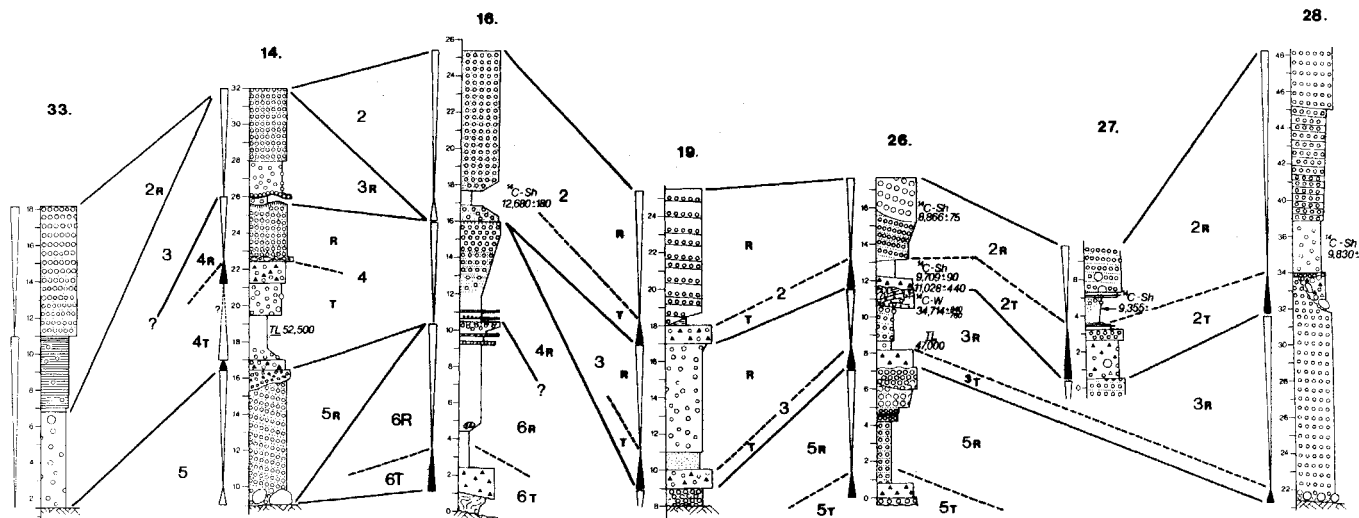
Figure 3a shows a graph of free against total D/L alloisoleucine from the sampled marine episodes. It shows initial racemization rates to be greater for the free acids, and as we are interested in high resolving power in relatively recent periods we use these as means of correlation in preference to total acids. Figure 3b shows a single mean free ratio for the marine component of each 'glacio-isostatic unit'. From the groups shown we identify five events at 17 sites, in which each event represents a short-lived phase of 'glacio-isostatic' submergence and emergence. These are formally ordered into events 2-6. The individual values from each of the marine units in fig. 3a have been ringed according to the event to which listing of the means in Fig. 3b suggests they belong. As racemization rates are temperature dependent we attempt to date these events by calibration of the amino acid data in Fig. 3a against <sup>14</sup>C and thermoluminescence dates<sup>4,9,10</sup> on the marine episodes. For the period 0-10,000 yr BP, <sup>14</sup>C dates on both shells and wood are used. For older periods, though many finite <sup>14</sup>C dates on shells are available, they are frequently regarded as unreliable, and thus the only <sup>14</sup>C date we use is a date of 34,000 yr BP on wood from Kap Ekholm.

To the list of glacio-isostatic events 2-6 in Fig. 3b we must add the youngest, glacial event 1. This is the glacial maximum attained during the Little Ice Age of recent centuries, when glaciers were substantially more advanced than any time during the past 10,000 yr.

The pattern of representation of events in Fig. 3b makes us confident that notwithstanding local erosion or non-deposition of evidence, Fig. 3b records all significant glacial episodes in Spitsbergen since at least event 4. We assess the timing and extent of the glacial episodes as follows:

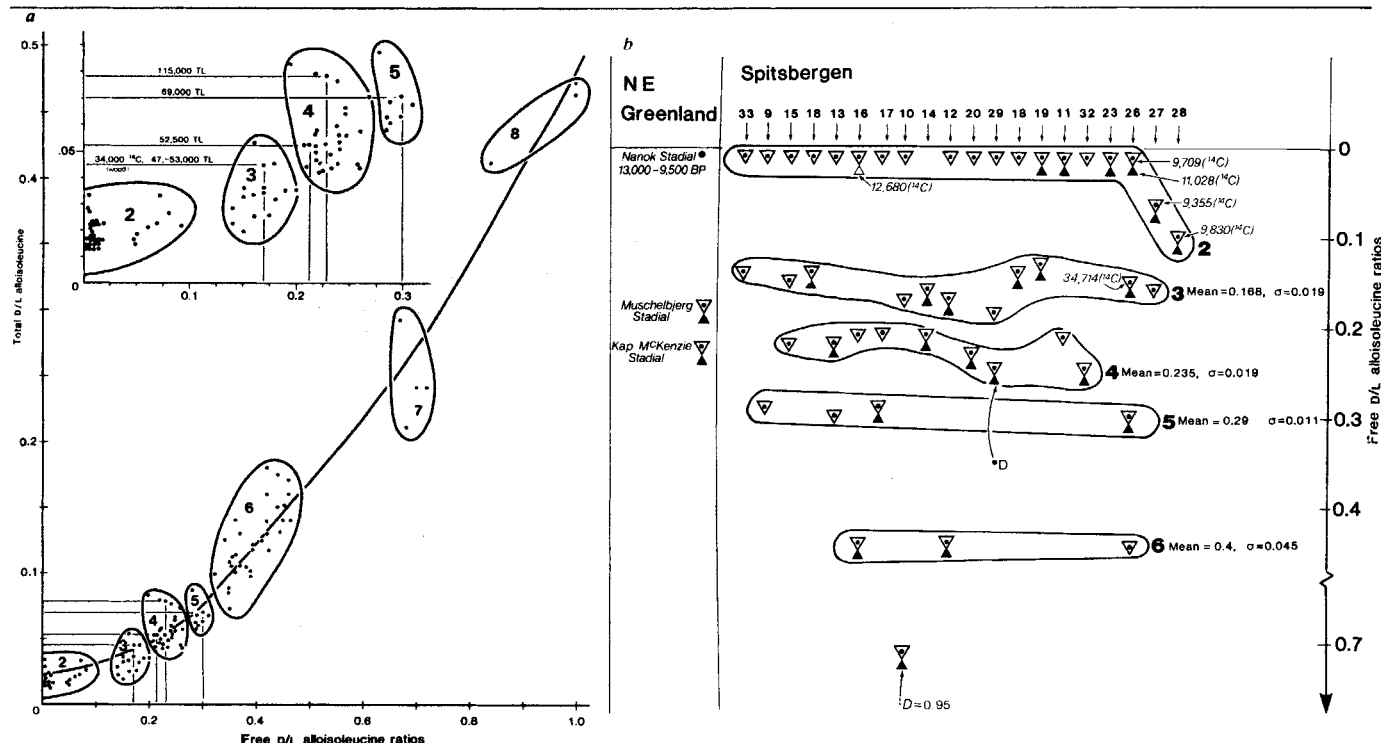
**Event 1** Little Ice Age Stadial: Advance of unknown magnitude (at least 12 km advance in some places).

**Event 2** Billefjord Stage Stadial: Major expansion of ice



**Fig. 2** Selected stratigraphic sequences from the Spitsbergen archipelago. Site locations are shown in Fig. 1a. The correlations from site to site show the glacio-isostatic units numbered as in Fig. 3b, and with regressive and transgressive components identified. Carbon-14 dates are on shell material (Sh) and driftwood (W).





**Fig. 3** *a*, Free D/L allosioleucine ratios plotted as points against total ratios for individual samples from marine units. The means from each unit are plotted in *b* and correlated as events 1-6. All the individual samples belonging to a single event are then grouped (2-8). The inset shows the relationship between amino acid ratios and thermoluminescent (TL) and  $^{14}\text{C}$  dates. (Dates are from refs 4, 9 and H. Pryczynska, personal communication.) *b*, Mean free D/L allosioleucine ratios plotted for the marine component of each glacio-isostatic unit at each site. The groupings 2-6 are suggested correlations for Spitsbergen. Data from north-west Greenland<sup>11</sup> are also plotted. Triangles indicate coarsening/fining sense of component sequences; ▲, 'glacial' components; △, marine components (see text). Events 7-8 are too poorly represented to allow conclusions to be drawn.

domes in east. Advances of several kilometres beyond modern margins in the west at some time between 12,000 and 10,000 yr BP. Eastern ice domes grew after ~40,000 yr BP and disappeared by 9,800 yr BP.

**Event 3** Stadial (mean free D/L ratio = 0.168,  $\sigma = 0.019$ ): More extensive advance in west Spitsbergen than during event 2. End of episode at about 35,000-40,000 yr BP. Events in east unknown, but any major ice domes disappear by the end of the event.

**Event 4** Stadial (mean free D/L ratio = 0.235,  $\sigma = 0.019$ ): Major glacial advance in west Spitsbergen. Glaciers advance at least to the mouths of the major west coast fjords. End of the event at some time between 50,000 and 115,000 yr BP. As an hypothesis, we suggest a median date of 75,000-80,000 yr for the end of this event.

**Event 5** Stadial (mean free D/L ratio = 0.29,  $\sigma = 0.011$ ): Glacial event ends at some time before 70,000 yr BP. This could be an early Weichselian event or pre-date the last interglacial.

**Event 6** Stadial (mean free D/L ratio = 0.4,  $\sigma = 0.045$ ): Glacial event before 115,000 yr BP. We suggest that it predates the last interglacial.

## Greenland correlations and implications for Arctic Ocean ice shelves

A mere 800 km to the west of Spitsbergen, on the north-west coast of Greenland, Hjort<sup>11</sup> has identified three Weichselian phases during which the eastern margin of the ice sheet underwent a major advance. The latest phase (Nanok Stadial) peaked between 13,000 and 9,500 yr BP, the time equivalent of the Billefjord Stadial in Spitsbergen. Two earlier episodes were both followed by marine transgressions from which amino acid analyses were made on mollusc faunas. Free D/L allosioleucine ratios (Fig. 3*b*) from these two transgressions show very similar values to those of events 3 and 4 in Spitsbergen. If we assume the transgressions to have been immediately post-glacial isostatic inundations of short duration (Fig. 1*b*), and in view of the

similar climate of the two areas, a time correlation is suggested between events 3 and 4 in Spitsbergen and the Muschelbjerg and Kap McKenzie Stadials in north-east Greenland, estimated by Hjort to be of mid-Weichselian and early Weichselian age respectively. All three Greenland Weichselian stadials have the same relative magnitudes of glacial extension as events 2, 3 and 4 in Spitsbergen (Fig. 4).

It has been suggested that the Arctic Ocean and Norwegian Sea were covered by a complex of thick ice shelves during the late Weichselian global glacial maximum (isotope stage 2); fed by glaciers flowing off the adjacent land areas<sup>12</sup>. However, the few Arctic areas where a Weichselian glacial stratigraphy is well developed (East Baffin Island<sup>13</sup>, Ellesmere Island<sup>14</sup>, East Greenland<sup>11</sup> and Spitsbergen) all show evidence of broad coastal zones occupied by raised beaches of mid-Weichselian age, demonstrating that these coastal areas have not been subsequently crossed by major glaciers, and all fail to show evidence of major late Weichselian glacial growth. On this evidence, the theory seems implausible, and increasingly more complex hypotheses must be used to sustain it.

## Weichselian glacio-climatic environments in the Atlantic sector

We suggest that the relative magnitudes and timing of glacial events in high Arctic latitudes of the Atlantic sector of the Northern Hemisphere are quite different from those in middle latitudes at the southern margins of the great Weichselian ice sheets<sup>15,16</sup>. They are compared in Fig. 4, together with contemporary movements of water masses in the Atlantic and Norwegian Sea<sup>17-19</sup>. The major stadial events in middle latitudes seem to coincide with major southerly expansions of the oceanic polar front in the Atlantic, the greatest glacial expansion being at 18,000-20,000 yr BP. In contrast, the greatest Weichselian glacier advance in the Arctic (stadial 4 in Spitsbergen and Kap McKenzie Stadial in Greenland) we suggest to have been in early Weichselian times, before to 75,000 yr BP, at which time there is no evidence of a major glacial event in Northern Europe or America, but rather of boreal forest, during a period in





into the Arctic than occurred in subsequent interstadials, though all shared the basic circulation pattern in Fig. 5b. This was also the period of greatest Weichselian accumulation rates of snow and ice on the North Greenland ice sheet<sup>24</sup>. We suggest that the Billefjord Stadial reflected an increased moisture flux into the Arctic some time after 15,000 yr BP and before 11,000 yr BP when the circulation in Fig. 5b broke down in late Weich-

selian time, and when this section was still relatively cold. It is not necessarily genetically related to the glacial episode which occurred between 11,000 and 10,000 yr BP in middle latitudes.

This work was supported by the Natural Environment Research Council and the NSF. We thank Mike Paul, Caroline Eyles, John Young, Raymond Jowett, Kari Aasgaard, Marcus Boddy and Jim Clark for valuable assistance in the field.

Received 22 March; accepted 13 May 1982.

1. Feyling-Hansen, R. W. *Norsk Polarinst. Medd.* **93**, 1–5 (1965).
2. Salvigsen, O. *Norsk Polarinst. Arb.* 217–228 (1981).
3. Salvigsen, O. *Geogr. Ann.* **63A**, 287–291 (1981).
4. Boulton, G. S. *Boreas* **8**, 31–57 (1979).
5. Boulton, G. S. & Rhodes, M. *Geol. Mag.* **111**, 481–500 (1974).
6. Schytt, V., Hoppe, G., Blake, W. & Grosswald, M. G. *Int. Ass. Sci. Hydrol.* **79**, 207–216 (1968).
7. Punning, J. M., Rajamae, R., Joers, R. & Putnik, H. *Azvestia Acad. Nauk. Estonsk., SSR. Geologya* **29**, 34–46 (1977).
8. Blake, W. in *Geology of the Arctic* (ed. Raasch, O.) 133–145 (University of Toronto Press, 1961).
9. Troitsky, L., Punning, J. M., Hütt, G. & Rajamae, R. *Boreas* **8**, 401–407 (1979).
10. Ruddiman, W. F. & McIntyre, A. *Science* **204**, 173–175 (1979).

11. Hjort, C. *Boreas* **10**, 259–274 (1981).
12. Hughes, T., Denton, G. H. & Grosswald, M. G. *Nature* **266**, 596 (1977).
13. Andrews, J. T. in *Quaternary Glaciations in the Northern Hemisphere* (Project 73/1/24, IGCP Rep 4, 1976).
14. England, R. & Bradley, R. S. *Science* **200**, 265 (1978).
15. Lundqvist, J. *Boreas* **10**, 395–402 (1981).
16. Dreimanis, A. & Karrow, P. *Int. Geol. Congr. Montreal, Sect. 8*, 13–24 (1972).
17. Kellogg, T. B. *Mem. geol. Soc. Am.* **145**, 77–110 (1976).
18. Ruddiman, W. F. & McIntyre, A. *Mem. geol. Soc. Am.* **145**, 11–146 (1976).
19. Ruddiman, W. F. *Bull. geol. Soc. Am.* **88**, 1813–1827 (1977).
20. Kellogg, T. B. *Boreas* **9**, 115–137 (1980).
21. Wright, P. B. & Lamb, H. H. *Climat. Res. Unit. Res. Publ.* **2**, 104–108 (1973).
22. CLIMAP *Science* **191**, 1131–1137 (1976).
23. Lamb, H. H. *Climate: Present, Past and Future*, Vol. 2 (Methuen, London, 1977).
24. Andrews, J. T., Funder, S., Hjort, C. & Imbrie, J. *Geology* **2**, 355–358 (1974).

# The N-terminal arms of $\lambda$ repressor wrap around the operator DNA

Carl O. Pabo\*, William Krovatin†, Andrea Jeffrey\* & Robert T. Sauer†

\* Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

† Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

*From differences in the ability of fragments of the  $\lambda$  repressor to protect operator bases from chemical modification we suggest that the first few N-terminal residues of the  $\lambda$  repressor form an extended arm that reaches around the back of the DNA helix when repressor binds to the operator.*

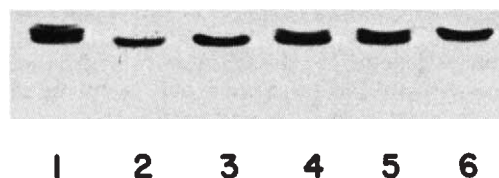
THE N-terminal, DNA-binding domain of the  $\lambda$  repressor (residues 1–92) protects the same operator guanines from chemical modification as intact repressor<sup>1</sup>. Crystallographic studies show that this domain contains five  $\alpha$ -helices and an extended N-terminal arm, and model-building studies indicate that two symmetrically related N-terminal domains bind to each 17-base pair operator site<sup>2</sup>. These studies also suggest that  $\alpha$ -helices from each domain contact the major groove near the outer regions of an operator site and that the N-terminal arms are near the centre of the operator site<sup>2</sup>. Here we present biochemical evidence that indicates that the N-terminal arms reach around the operator helix and contact the major groove on the 'back' of DNA.

## Tryptic cleavage of residues 1–3 of the N-terminal domain

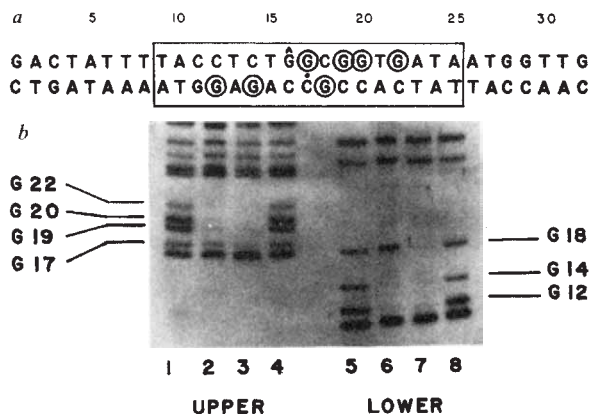
Digestion of repressor with papain generates a fragment (residues 1–92) that comprises the N-terminal domain<sup>3</sup>. When this fragment (fragment d) is treated with trypsin and the products analysed by native gel electrophoresis (Fig. 1), the band corresponding to fragment d disappears, and a new band appears. This new fragment was purified (see Fig. 1 legend), and was subjected to sequential Edman degradation. The new N-terminal sequence (NH<sub>2</sub>-Lys-Lys-Pro-Leu...) showed that the first three residues of repressor (Ser-Thr-Lys) had been removed<sup>4</sup>. Amino acid composition data (not shown) and the behaviour of the new fragment on native gels (Fig. 1) indicated that trypsin had not altered the C-terminal sequence of fragment d. We conclude that trypsin cleaves the Lys 3-Lys 4 peptide bond of fragment d, and refer to the trypsin modified N-terminal fragment as fragment d<sub>4–92</sub>.

## Interaction of fragment d<sub>4–92</sub> with operator DNA

We have shown previously that fragment d and intact  $\lambda$  repressor produce identical patterns of methylation protection and enhancement at the  $\lambda$  operator site O<sub>R</sub>1 (ref. 1). Here, we



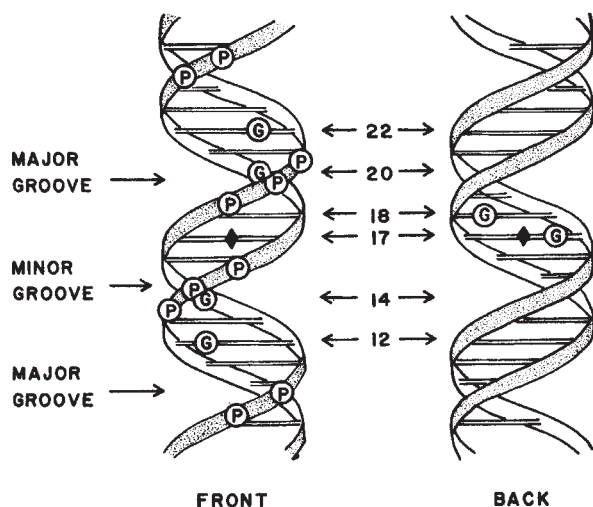
**Fig. 1** Analysis of fragment d and modified fragment d by native gel electrophoresis. The electrophoresis buffer consisted of 30 mM sodium acetate-acetic acid (pH 4.5), and the negative electrode was on the bottom of the slab gel. Lanes 2 and 3 show intact fragment d. Lanes 4–6 show the products of tryptic digestion of fragment d. Digestions were performed with fragment d at 1 mg ml<sup>-1</sup> in 30 mM bis-Tris propane-HCl (pH 7), 1 mM Na<sub>3</sub>N<sub>3</sub> for 2 h at 0°C. Digestion was terminated by addition of phenylmethylsulphonyl fluoride. Enzyme: substrate ratios (w/w) were 1:1,000 (lane 4), 1:500 (lane 5) and 1:50 (lane 6). Lane 1 shows fragment d that has been treated with an approximately equimolar amount of acetic anhydride. The lowest band is unmodified fragment d. Singly and doubly acetylated fragments have lost one or two positive charges at this pH and migrate more slowly. Comparison of lanes 4–6 with lane 1 shows that the product of tryptic digestion of fragment d migrates at the same rate as singly acetylated fragment d and thus appears to have lost one positive charge. This is expected if the tryptic fragment has lost Ser 1-Thr 2-Lys 3. Any other tryptic cleavage of fragment d, or any additional cleavage of fragment d<sub>4–92</sub> would generate species with a net charge different from that of fragment d<sub>4–92</sub> (ref. 4). No such species were detected. Preparative tryptic digestions of fragment d were monitored by native gel electrophoresis to ensure complete digestion. Following digestion, the sample was dialysed into 30 mM sodium acetate-acetic acid (pH 5) and was loaded onto a column of sulphopropyl-Sephadex (C-25). The column was eluted with a gradient of increasing NaCl concentration, and fragment d<sub>4–92</sub> was located by absorbance at 280 nm. The pooled sample was concentrated by ultrafiltration and was further purified by chromatography on a Sephadex G-75 (superfine) column.



**Fig. 2** *a*,  $\lambda$  operator site  $O_{R1}$ . Bases within the boxed region comprise a repressor binding site<sup>6,11</sup>. Repressor suppresses (circled Gs) or enhances (G) methylation of the indicated guanines<sup>6</sup>. *b*, Effects of fragments d and  $d_{4-92}$  on methylation of  $O_{R1}$ . Reactions were performed at 0°C as described in ref. 1. Lanes 1–4 and 5–8 show the upper and lower strands of  $O_{R1}$  respectively. Additions were as follows: lanes 1, 4, 5 and 8, no protein; lanes 2 and 6,  $2 \times 10^{-4}$  M fragment  $d_{4-92}$ ; lanes 3 and 7,  $2 \times 10^{-4}$  M fragment d.

compare the operator binding of fragments d and  $d_{4-92}$ . Figure 2*b* shows methylation of  $O_{R1}$  in the presence of fragment d (lanes 3, 7) or fragment  $d_{4-92}$  (lanes 2, 6). Control lanes (1, 4, 5, 8) show methylation in the absence of protein. Comparison of lanes 6 and 7 reveals a striking difference between the protection patterns produced by the two repressor fragments. Fragment d protects guanines 12, 14 and 18 on the lower strand of  $O_{R1}$  (lane 7); fragment  $d_{4-92}$  protects only guanines 12 and 14 (lane 6). The fragments also produce different patterns of methylation protection and enhancement on the upper strand of  $O_{R1}$ . Fragment d protects guanines 17, 19, 20 and 22 (lane 3); the shortened fragment protects guanines 20 and 22 and partially protects guanine 19 (lane 2). Also, the methylation of guanine 16 is enhanced by fragment d but not by the shorter fragment.

Protection experiments like those shown in Fig. 2*b* were repeated at a series of fragment concentrations (data not shown). Fragment d gave half-maximal protection of  $O_{R1}$  at a concentration of  $2 \times 10^{-8}$  M and fragment  $d_{4-92}$  gave half-maximal protection at a concentration of  $10^{-6}$  M. Since the

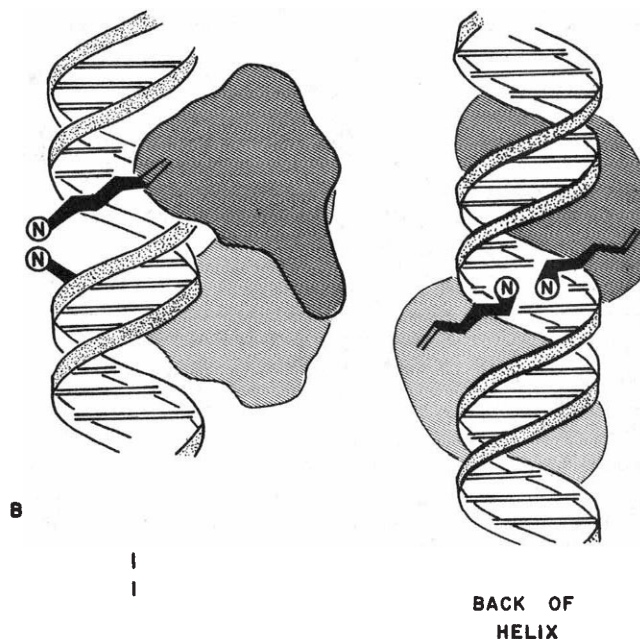


**Fig. 3** Front and back views of  $O_{R1}$ . Positions marked by G represent guanines whose methylation interferes with repressor binding<sup>6</sup>. Positions marked by P represent phosphates whose ethylation interferes with repressor binding<sup>6</sup>.  $G^{19}$  is not shown in these diagrams because it would be visible from the side of the helix. In each view, the approximate 2-fold axis is perpendicular to the plane of the page and is marked with a diamond.

concentration of operator DNA was low in these experiments, the concentration of protein required for half-maximal protection corresponds to the apparent equilibrium dissociation constant for the fragment-operator complex. Converting these dissociation constants to binding energies indicates that removing the three N-terminal residues of fragment d reduces its affinity for  $O_{R1}$  by about 2 kcal mol<sup>-1</sup>.

### Residues 1–3 contact guanines 17 and 18

It is thought that proteins protect guanines from methylation by contacting or blocking the reactive N-position<sup>5</sup>. This steric interpretation seems reasonable in the  $\lambda$  repressor case, as prior methylation of any protected guanine reduces the affinity of repressor for the modified operator<sup>6</sup>. Our results thus suggest that fragment d makes operator contacts that fragment  $d_{4-92}$  fails to make. We believe that fragments d and  $d_{4-92}$  have essentially identical structures and infer that the differences in



**Fig. 4** Sketches showing how the N-terminal arms might reach around the DNA to contact sites in the major groove on the back of the operator. The position and shape of the globular portion of the N-terminal domain are taken from ref. 2. Structural details have been omitted for simplicity.

their methylation protection patterns result from contacts normally provided by residues 1–3 of repressor. Residues 1–3 are disordered in the electron density map of fragment d (ref. 2), and thus it is unlikely that they stabilize the remaining protein structure, so removing these residues should not perturb the basic structure of the globular N-terminal domain. This conclusion is supported by NMR studies of fragments d and  $d_{4-92}$  (M. Weiss and D. Patel, personal communication).

### Residues 1–3 contact bases on the back of the operator helix

Figure 3 shows  $O_{R1}$  represented as B-DNA, and indicates the guanine N-7 and phosphate positions that apparently interact with repressor. We represent  $O_{R1}$  as B-DNA since experiments indicate that repressor does not significantly change the pitch of the DNA helix when it binds<sup>7</sup>. The left half of Fig. 3 shows that four of the guanine 'contacts' and all of the phosphate 'contacts' may be seen from one side of the DNA helix, which we refer to as the 'front' of the operator. Model building studies suggest that  $\alpha$ -helices of repressor contact the major groove on this side of the DNA molecule<sup>2</sup>.

The right half of Fig. 3 shows that two of the guanine contacts (guanines 17 and 18) are located on the back of the operator



helix. These bases on the back are precisely those protected by fragment d but not by fragment d<sub>4-92</sub>. We conclude that residues 1-3 of repressor must somehow reach around the DNA helix to make these contacts.

### A model for the repressor-operator interaction

In Fig. 4, we indicate schematically how the N-terminal arms of repressor might wrap around the DNA helix. In these sketches, the positions of the globular portion of the N-terminal domains have been taken from model-building studies based on the crystal structure<sup>2</sup>. The positions of the arms have been inferred from the biochemical results presented here. The repressor arms have been placed in the major groove as the reactive N-7 position of the contacted guanines is exposed in this groove. If the arms are built in an extended conformation, residues 1-3 can be positioned close to the N-7 positions of guanines 16-19, where fragments d and d<sub>4-92</sub> have different effects on methylation. Although specific chemical contacts between the arm and the DNA are not defined by these studies, Ser 1 and Thr 2 might form hydrogen bonds to the bases, and Lys 3, Lys 4 and Lys 5 might form salt bridges with the phosphates. We do not know if contacts made by residues of the arm are sequence specific, however, the *v*<sub>3</sub> operator constitutive mutation in O<sub>R</sub>1 changes guanine 18 to thymine<sup>8</sup>, providing further evidence that repressor contacts the centre of the operator site and suggesting that some of the contacts made by the arm may be sequence specific.

### Discussion

A flexible or extended structure could play several special roles in protein-DNA interactions. Such structures would allow a protein to contact both sides of the helix without creating a large kinetic barrier to association. We believe that the  $\lambda$  operator site is completely encircled by the bound repressor.

Received 2 February; accepted 13 May 1982.

1. Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M. & Backman, K. C. *Nature* **279**, 396-400 (1979).
2. Pabo, C. O. & Lewis, M. *Nature* **298**, 443-447 (1982).
3. Pabo, C. O., Sauer, R. T., Sturtevant, J. M. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1608-1612 (1979).
4. Sauer, R. T. & Andereg, R. *Biochemistry* **17**, 1092-1100 (1978).

As a preformed  $\lambda$  repressor dimer binds to the operator site<sup>9</sup>, a completely rigid repressor protein would have to be threaded onto a free end of the DNA—the use of flexible arms avoids this problem. The arms can wrap around the helix after the globular portion of the repressor has bound to the front of the operator helix. However, the flexibility of the arm may limit the amount of binding energy it can provide. If  $\lambda$  repressor's N-terminal arms are disordered in solution but adopt a specific conformation (or a limited number of conformations) in the complex with DNA, there will be a loss of conformational entropy as repressor binds. This implies that contacts between the arms and the DNA cannot provide as much free energy of binding per contact as a rigid region could provide<sup>10</sup>.

A flexible region might also allow a DNA-binding protein to adjust to the asymmetric environment that may exist at the centre of an operator site. For example, the  $\lambda$  operator sites contain 17 base pairs and the axis of approximate 2-fold symmetry intersects the central base pair. The central base pair cannot obey the 2-fold symmetry axis (as this would interchange a purine and a pyrimidine) and thus it would be impossible for a repressor dimer to contact this base pair in a symmetric manner. Flexibility of the arms might allow one arm to move aside while the other contacts the central base pair. An asymmetric arrangement at the centre could also affect interactions at nearby (symmetric) bases.  $\lambda$  repressor's methylation protection pattern at O<sub>R</sub>1 is consistent with this possibility: G<sup>16</sup> and G<sup>18</sup> are strictly related by the 2-fold axis, but bound repressor enhances the methylation of G<sup>16</sup> and suppresses methylation of G<sup>18</sup>.

This work was supported by NIH grants AI-16892 to R.T.S. and NIH grant GM-22526 to M. Ptashne. We thank A.D. Johnson for helpful discussions and M. Weiss and D. Patel for communication of unpublished data. C.O.P. was supported by a postdoctoral fellowship from the Jane Coffin Childs Foundation for Medical Research.

5. Gilbert, W., Maxam, A. & Mirzabekov, A. D. *Alfred Benson Symp.* **9**, 139-148 (1976).
6. Johnson, A. D. *J. molec. Biol.* (in the press).
7. Maniatis, T. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1531-1535 (1973).
8. Flashman, S. M. thesis, Harvard Univ. (1976).
9. Sauer, R. T. thesis, Harvard Univ. (1979).
10. Huber, R. *Trends biochem. Sci.* **4**, 271-276 (1979).
11. Ptashne, M. *et al. Cell* **19**, 1-11 (1980).

# The operator-binding domain of $\lambda$ repressor: structure and DNA recognition

Carl O. Pabo & Mitchell Lewis

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

*The structure of the operator-binding domain of the  $\lambda$  repressor has been determined at 3.2 Å resolution. This domain contains an extended N-terminal arm and five  $\alpha$ -helices. Model-building studies of the repressor-operator complex suggest that  $\alpha$ -helices, especially the N-terminal parts of these helices, may provide a useful surface for protein-DNA interactions.*

THE interaction of proteins with specific sites on double-stranded DNA is essential for many aspects of cellular control. To understand these control systems, one must determine how proteins recognize specific sites on double-stranded DNA, how the bound proteins interact with other regulatory proteins, and how the binding may affect the conformation and accessibility of other sites on the DNA. To address these general questions, we have undertaken a crystallographic study of the  $\lambda$  repressor and its operator.

The  $\lambda$  repressor stimulates transcription of its own gene, and turns off transcription of the other phage genes by binding to two operator regions in the phage DNA<sup>1,2</sup>. Repressor dimers

bind to three sites in each region. Each repressor-binding site is 17 base pairs long, and the sequence of each site has an approximate 2-fold rotational symmetry. The  $\lambda$  repressor, which has 236 amino acid residues<sup>3</sup>, contains two domains<sup>4</sup>. Cleavage of the repressor with the protease papain yields a 92 residue N-terminal fragment which binds specifically to the  $\lambda$  operators, apparently making all the operator contacts that the complete repressor makes<sup>5,6</sup>. Like the complete repressor, this fragment can mediate positive and negative control of transcription<sup>7</sup>. The C-terminal domain allows repressor to form stable dimers that have a higher operator-binding affinity than the isolated N-terminal fragment.



**Table 1** Data collection (a) and phasing (b) statistics

a

| Data set   | No. crystals | No. reflections measured | Unique reflections | $R_{\text{sym}}$ |
|------------|--------------|--------------------------|--------------------|------------------|
| Native     | 5            | 14,276                   | 5,984              | 0.051            |
| Derivative | 8            | 14,096                   | 11,317             | 0.067            |

b

| Group | No. of reflections | Resolution limits (Å) | $m$   | Isomorphous $f_{\text{r.m.s.}}/E_{\text{r.m.s.}}$ | $R_K$ | Average $R$ | $\Delta\Phi$ |
|-------|--------------------|-----------------------|-------|---------------------------------------------------|-------|-------------|--------------|
| 1     | 295                | 9.1–50.0              | 0.681 | 2.99                                              | 0.098 | 0.565       | 78.97        |
| 2     | 537                | 6.4–9.1               | 0.711 | 3.05                                              | 0.117 | 0.175       | 54.27        |
| 3     | 656                | 5.2–6.4               | 0.700 | 3.41                                              | 0.118 | 0.179       | 56.47        |
| 4     | 761                | 4.5–5.2               | 0.682 | 2.78                                              | 0.115 | 0.177       | 56.29        |
| 5     | 875                | 4.0–4.5               | 0.681 | 2.21                                              | 0.112 | 0.184       | 56.35        |
| 6     | 983                | 3.7–4.0               | 0.651 | 2.29                                              | 0.099 | 0.197       | 62.76        |
| 7     | 1,012              | 3.4–3.7               | 0.561 | 2.41                                              | 0.095 | 0.219       | 66.30        |
| 8     | 848                | 3.2–3.4               | 0.531 | 2.40                                              | 0.112 | 0.221       | 68.35        |
| Total | 5,967              | 3.2–50.0              | 0.630 | 2.54                                              | 0.106 | 0.220       | 61.74        |

a, Crystals were transferred to sodium citrate (pH 5.9; 95% saturated), and a platinum derivative was prepared by soaking the crystals in 1 mM  $\text{K}_2\text{PtCl}_4$  for 24 h. Data were collected on a Nicolet P3F diffractometer using Ni-filtered radiation, and the crystals were cooled to 0 °C during data collection. Anomalous data were collected from the platinum derivative. The mean fractional isomorphous difference was 0.245.

$$R = \frac{\sum |I - \bar{I}|}{\sum \bar{I}}$$

b, Statistics from the SIR phasing are summarized in the three columns labelled 'Isomorphous'. The refined platinum-binding sites formed an almost perfect equilateral triangle, and these sites were used to define the location and orientation of the local 3-fold axis. The SIR map was averaged around this axis, and the averaged map was used to define a preliminary molecular envelope. At the start of each cycle, a new map was calculated. Points within the envelope were averaged to impose the 3-fold symmetry, and points outside the envelope were set to zero. This map was inverted to give  $F_c$  and  $\alpha_c$ . The next cycle used  $2F_0 - F_c$  and  $\alpha_c$ . As the map improved, a new molecular envelope was drawn, and the averaging was repeated until the solution converged. The last columns of this table give statistics from the averaging. The  $R$  factor compares the observed structure factors with structure factors calculated from the symmetrized electron density map. The last column gives the average absolute difference between the SIR phases and the final phase set.  $m$ , mean figure of merit.  $\Delta\Phi$ , average absolute difference in degrees between SIR phases and final phase set.  $F_c$  was calculated from the symmetrized electron density map.

$$f_{\text{r.m.s.}} = \left( \frac{\sum f_H^2}{n} \right)^{1/2}$$

$$E_{\text{r.m.s.}} = \left( \frac{\sum (F_{\text{PH}} - |F_P + f_H|)^2}{n} \right)^{1/2}$$

$$R_K = \frac{\sum |F_{\text{PH}}^0| - |F_P^c + f_H|}{\sum |F_{\text{PH}}^0|}$$

$$R = \frac{\sum |F_0 - F_c|}{\sum |F_0|}$$

Here we report the structure of the N-terminal domain of  $\lambda$  repressor at 3.2 Å resolution. We suggest a model for the repressor-operator interaction and discuss features of this model that may be relevant to other protein-DNA interactions.

## Structure determination

The N-terminal fragment of repressor was prepared as described elsewhere<sup>7</sup>. The purified fragment was dialysed against a buffer containing 30 mM bis-Tris propane-HCl (pH 7.0) and 1 mM sodium azide. The protein was concentrated to 40 mg ml<sup>-1</sup>, and the solution was centrifuged. Crystals were grown by the hanging drop vapour diffusion method, using a mixture of ammonium phosphate (pH 5.9, 48% saturated) and sodium chloride (20% saturated) as a precipitant at 4 °C. The crystals, which grew in about 2 months, developed as trigonal or hexagonal plates that were up to 1 mm across and 0.2 mm thick.

Precession photographs revealed that the crystals formed in space group  $P3_121$  (or  $P3_221$ ) and had cell dimensions of  $a = b = 65$  Å,  $c = 149.6$  Å. The volume of the unit cell indicated that there were several fragments in the asymmetric unit, and an unusual pattern of weak reflections allowed us to determine how these fragments were arranged. Reflections tended to be weak if both  $h$  and  $k$  were even and if  $-h + k + 1 = 3n$ . This pattern of weak reflections implied that the molecules were arranged in layers in the unit cell and that each layer had a local 3-fold axis which was approximately parallel to the  $c$  axis<sup>7</sup>. Intensity data from native crystals and from a platinum derivative were measured on a diffractometer. Statistics are summarized in Table 1a, and some details of data collection are given in the legend.

The platinum positions were located by inspection of difference Pattersons. These sites were refined, and SIR phases

were calculated (see Table 1b). The SIR map was noisy, but it allowed us to locate the molecules and indicated that the correct space group was  $P3_221$ . As there were three copies of the protein in the asymmetric unit, non-crystallographic symmetry averaging was used to improve the quality of the electron density map<sup>8</sup>. Seven cycles of averaging were used to obtain a refined set of phases. Phasing statistics are given in Table 1b, and some details of the averaging are given in the legend. Figure 1 shows sections from a SIR map and from a map calculated with the refined phases.

Initial inspection of the new electron density map revealed several  $\alpha$ -helices. The polypeptide chain was traced on a minimap (4 Å per cm), and a detailed atomic model was built on a computer graphics system (using programs written by R. Diamond<sup>9</sup> and modified and implemented at Harvard by R. C. Ladner). There was strong electron density for most of the side chains, and the interpretation of the map was fully consistent with the known amino acid sequence<sup>3</sup>. However, there was no density for Ser 1-Thr 2-Lys 3, suggesting that these residues are disordered in the crystals or have different conformations in the three molecules that were averaged.

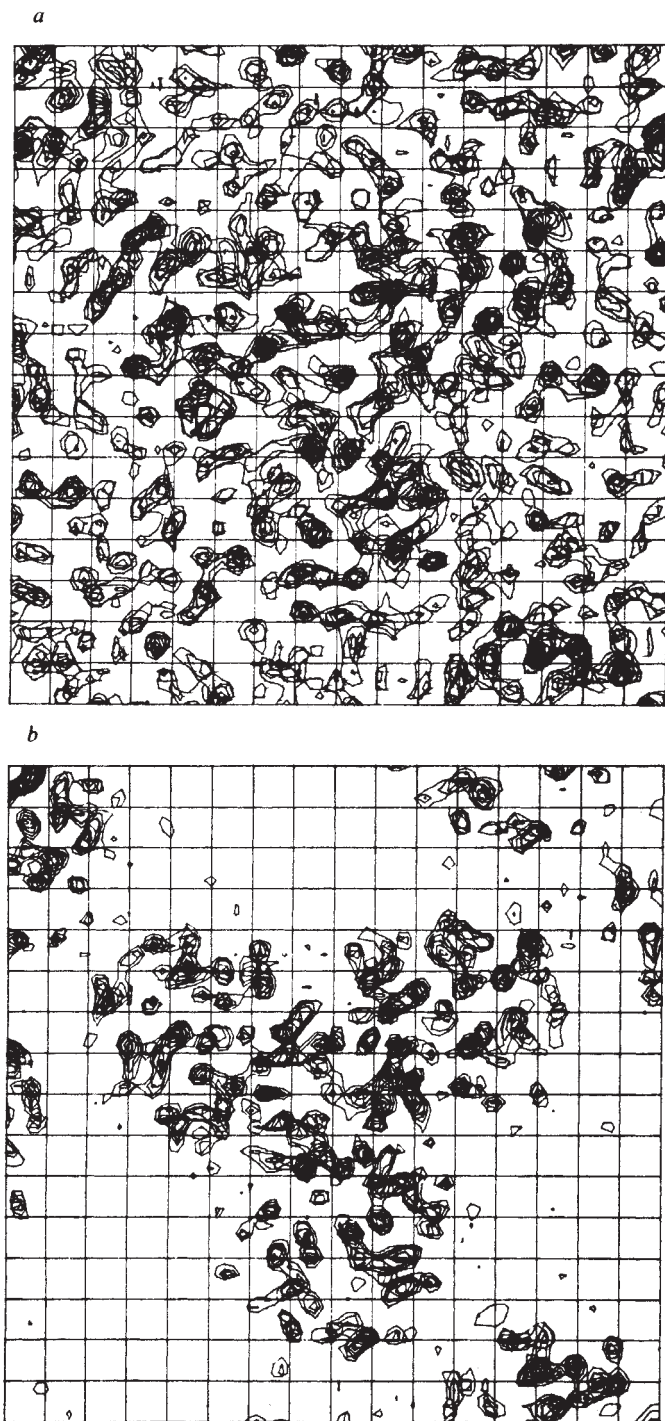
## Conformation of the N-terminal domain

The N-terminal domain contains an extended 'arm' and five  $\alpha$ -helices (residues 9–23, 33–39, 44–52, 61–69, and 79–92). The conformation of this domain is illustrated in Fig. 2. Residues 4–8 form the extended arm, which runs away from the surface of the monomer and packs against another molecule (which is related by the local 3-fold axis). Helix 1 starts with residue 9 and ends just before Lys 24, Lys 25, and Lys 26. Helix 2 includes residues 33–39. The chain makes a turn with the sequence Met 40-Gly 41-Met 42-Gly 43, and the platinum binds between the sulphur atoms of Met 40 and Met 42. The

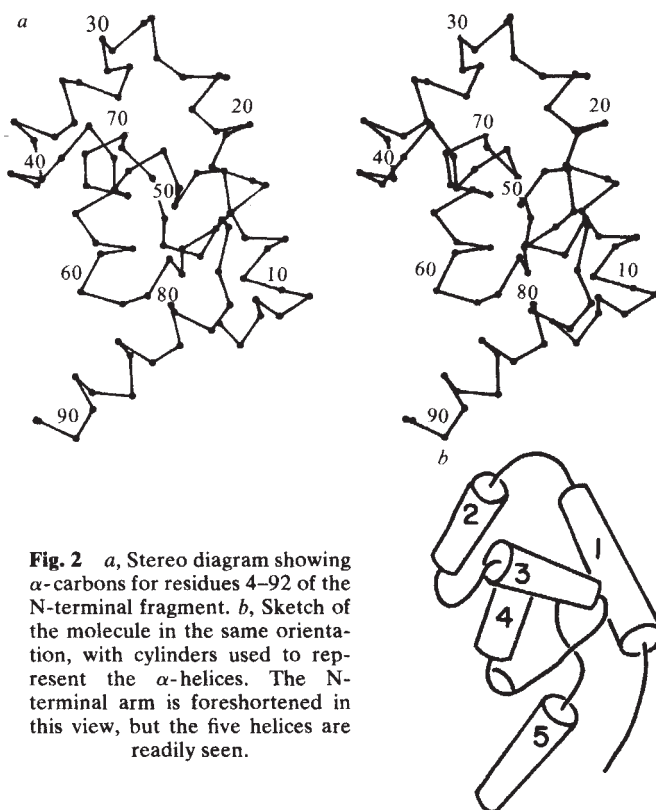
N-terminal part of helix 3 forms a protruding ridge on the surface of the molecule, but the C-terminus of this helix abuts helix 1 and is less exposed. After helix 3, an extended section of chain runs across a surface of the molecule and connects to helix 4. Helix 4 and the loop following complete the enclosure of a well formed hydrophobic interior, which includes about a dozen side chains. Helix 5 extends off to one side of the molecule and makes few contacts with the rest of the monomer.

### Intermolecular contacts in the crystal

In the crystal, the N-terminal fragment forms a dimer that is stabilized by an interaction between helix 5 of one fragment and helix 5 of a second fragment. The helix-helix contact is



**Fig. 1** Electron density maps before and after phase averaging. *a*, Map calculated with SIR phases. Sections are perpendicular to the local 3-fold axis, which is in the centre of this map. Lines on the orthonormal grid mark 5 Å intervals. *b*, Corresponding section of the final electron density map.



**Fig. 2** *a*, Stereo diagram showing  $\alpha$ -carbons for residues 4–92 of the N-terminal fragment. *b*, Sketch of the molecule in the same orientation, with cylinders used to represent the  $\alpha$ -helices. The N-terminal arm is foreshortened in this view, but the five helices are readily seen.

extremely hydrophobic and involves the side chains of Ile 84, Tyr 85, Met 87, and Tyr 88. In addition, the C-terminal part of helix 5 makes a few contacts with the globular region of the 2-fold-related molecule. Figure 3 shows a sketch of the dimer. Based on arguments given in the next section, we suggest that this dimer is the operator-binding species.

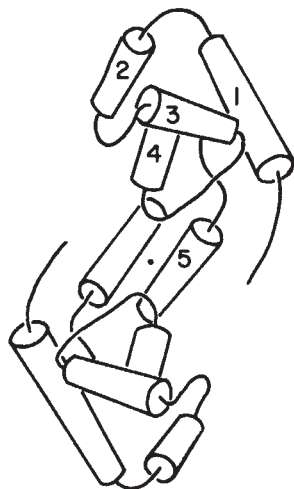
Three dimers cluster around the non-crystallographic symmetry axis to form a hexamer. This hexamer provides the basis for the crystal-packing arrangement, but we do not believe that it has any biological significance. The intermolecular contacts that stabilize the hexamer are less extensive than those that stabilize the dimer, and there is no evidence that hexamers have any biological function<sup>1</sup>. The extended N-terminal arm stabilizes the hexamer, because Lys 4, Lys 5, and Pro 6 pack against helix 5 of a 3-fold-related molecule. In addition, the side chains of Glu 10 from the three molecules cluster around this local symmetry axis, and the side chains of Glu 13 form a second cluster near this axis. This second cluster includes the side chains of Gln 9 and Arg 17, which are slightly further from the 3-fold axis.

### Interaction with the operator

To determine how repressor contacts the operator, we looked for features of the protein structure that were complementary, in shape and in bonding capability, to the structure of the  $\lambda$  operator site. Information from chemical and biochemical experiments provided constraints for our search. (1) The operator should be right-handed B-DNA, as repressor binding does not have a drastic effect on the helical pitch<sup>10</sup>. (2) Repressor binds primarily along one face of the double helix and contacts the DNA in the major groove<sup>11,12</sup>. (3) A repressor fragment that lacks the first three residues is unable to protect guanine residues on the 'back' of the operator site, and this indicates that the first residues wrap around the DNA<sup>13</sup>. (4) Both the base sequence of the operator sites and the results of chemical and enzymatic protection experiments indicate that the repressor-operator complex has approximate 2-fold rotational symmetry<sup>1,12</sup>. (5) Repressor forms dimers that bind to 17 base pair operator sites<sup>1,14</sup>.

Although the N-terminal fragment does not form a stable dimer in solution, two experiments suggest that it dimerizes as





**Fig. 3** Sketch of the N-terminal dimer, with cylinders used to represent helices. The orientation of the upper monomer is similar to the orientation in Fig. 2, and thus the N-terminal arm and helix 3 are closest to the viewer. The dot next to helix 5 represents the 2-fold axis, which is perpendicular to the page.

it binds to the operator<sup>4,7</sup>. The apparent operator-binding constant of the fragment is higher than one would expect for a monomer, but is readily explained if one assumes that the N-terminal fragment forms a weak dimer. Also, DNase protection experiments have failed to find protection of just one half of an operator site by the N-terminal fragment.<sup>6</sup>

To simplify our search for the proper complex, we assumed that the dimer observed in the protein crystal was the operator-binding species. The 2-fold axis of this dimer was superimposed on the approximate 2-fold axis of a  $\lambda$  operator site, and possible arrangements for the complex were evaluated on a computer graphics system. The fragment dimer was rotated about the 2-fold axis (and translated along this axis) until the fit appeared reasonable. Only the arrangement shown in Fig. 4 allowed the protein to make extensive contacts with the major groove of the operator site.

The proposed complex makes chemical as well as structural sense. A number of charged and polar groups are near the DNA. These include the side chains of Gln 33 and Glu 34, which are at the N-terminus of helix 2. The side chains of Gln 44 and Ser 45, which are at the N-terminus of helix 3, project directly into the major groove and probably hydrogen bond with the base pairs. Asn 52, at the C-terminus of helix 3, and both Asn 55 and Asn 58, which are in the loop after helix 3, also are very close to the DNA.

In the suggested complex, the extended N-terminal arms are set slightly to the side of the double helix and extend towards the back of the DNA. This is reasonable, as chemical protection experiments indicate that the first three residues of repressor wrap around the double helix and contact the back of the helix near the centre of the operator site<sup>13</sup>. Our model indicates that this arm could readily wrap around the helix, following the major groove, and reach around to the back of the DNA. Residues Ser 1, Thr 2, Lys 3, Lys 4 and Lys 5 could hydrogen bond to bases in the major groove or form salt bridges to the phosphates.

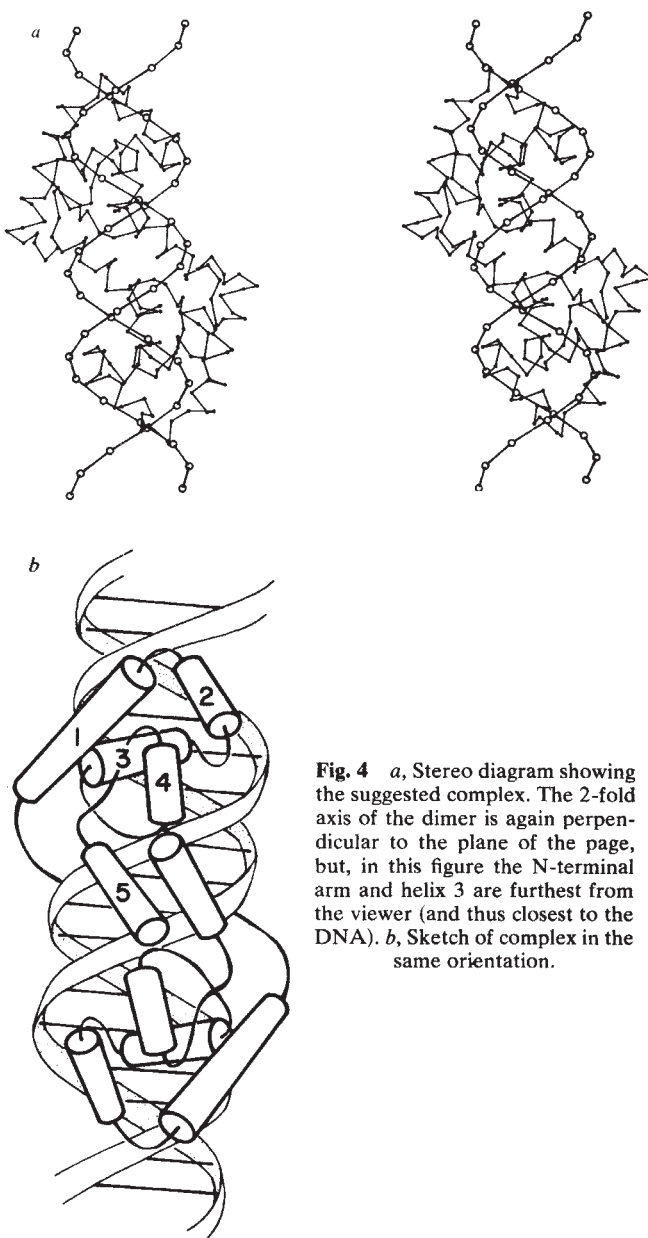
### Comparison with $\lambda$ cro

The  $\lambda$  cro protein binds to the same operator sites as the  $\lambda$  repressor, and the structure of cro has been determined recently<sup>15</sup>. The  $\lambda$  cro and the  $\lambda$  repressor share only limited sequence homology<sup>3,16,17</sup>, and the overall structures are quite different, since cro has three strands of  $\beta$ -sheet and three  $\alpha$ -helices. However, helices 2 and 3 of cro are clearly related to helices 2 and 3 of repressor<sup>17</sup>, and the suggested modes of DNA binding are basically similar.

Residues from helix 3 of repressor and from helix 3 of cro appear to contact bases in the major groove, and in each case, a Gln-Ser sequence at the N-terminus of this helix seems to contact the operator. However, the orientation of these helices in the suggested complexes is somewhat different. Helix 3 of cro is approximately parallel to the (local) direction of the major groove and the residues along one face of this  $\alpha$ -helix project

into the major groove. For repressor, the N-terminal portion of helix 3 is closer to the double-helical axis than is the C-terminal portion, and the N-terminal part of this helix seems to make most of the contacts with the major groove. To some extent, this may be seen in Fig. 4, which shows that the C-terminus of helix 3 is set to one side of the double-helical axis. Although it may not be obvious from Fig. 4, helix 3 also is inclined by 20° with respect to the plane of the page, and the C-terminus is closer to the viewer. In addition, we note that helix 3 of repressor is not precisely aligned with the major groove. If the double helical axis of the DNA is vertical, helix 3 of repressor will be inclined by about 16° with respect to a horizontal plane, whereas the major groove of B-DNA is inclined by about 32°.

In this context, two other differences between helix 3 of repressor and helix 3 of cro seem significant. (1) The middle and the C-terminal region of the repressor helix are significantly less polar than the corresponding regions of the cro helix. Thus two of the three polar residues in this part of the cro helix are aligned with nonpolar residues in the repressor helix. (2) The C-terminal part of the repressor helix does not protrude from the surface of the protein and could not fit into the major groove, even if the dimer contacts were altered. Any attempt



**Fig. 4** *a*, Stereo diagram showing the suggested complex. The 2-fold axis of the dimer is again perpendicular to the plane of the page, but, in this figure the N-terminal arm and helix 3 are furthest from the viewer (and thus closest to the DNA). *b*, Sketch of complex in the same orientation.



to move the C-terminal part of helix 3 much closer to the major groove would cause residues 54 and 55 to collide with the DNA.

## Repressor uses the N-terminal region of an $\alpha$ -helix to contact the DNA

Recent structural studies of  $\lambda$  cro and of CAP<sup>15,18</sup>, and a number of theoretical studies have suggested that  $\alpha$ -helices could fit into the major groove of B-DNA<sup>19-22</sup>. Our studies further support this idea. However, the role of the  $\alpha$ -helices in our model is somewhat different from that which has been emphasized in other studies. Previous studies have considered arrangements in which the axis of the  $\alpha$ -helix is parallel to path of the major groove and in which the centre of the  $\alpha$ -helix is closest to the double-helical axis of the DNA. Thus, one surface of the  $\alpha$ -helix fits completely into the major groove. Our model for the  $\lambda$  repressor-operator interaction indicates that the N-terminal part of helix 3 fits furthest into the major groove. Several general features of this arrangement should be noted: (1) Experiments with space-filling models suggest that residues on the N-terminal part of an  $\alpha$ -helix can contact three or four base pairs in the major groove. (2) The orientation of peptide dipoles in an  $\alpha$ -helix gives a partial positive charge at the N-terminus of the helix<sup>23</sup>. Placing this end near the DNA should allow a favourable electrostatic interaction. (3) The side chains of an  $\alpha$ -helix tend to point towards the N-terminus, and thus

they will be readily accessible for interactions with the DNA. The peptide hydrogens are also exposed on this end of the helix and could be used for hydrogen bonding.

We expect that other DNA-binding proteins will use the N-terminal region of an  $\alpha$ -helix for DNA binding. Actually, it is possible that the CAP protein uses the N-terminal parts of  $\alpha$ -helices to contact its site. It has been suggested that the CAP protein binds to left-handed B-DNA and that the F helices fit into the major grooves of this DNA<sup>18</sup>. However, coordinates taken from Fig. 6 of ref. 18 indicate that the N-terminal region of each F helix in the CAP dimer could fit into the major groove of right-handed B-DNA. This would allow helix F of CAP to be used in an arrangement that is quite similar to the one which we propose for helix 3 of the  $\lambda$  repressor. In addition, proteins (such as restriction enzymes) that contact short, symmetrical sites might use the N-terminal regions of  $\alpha$ -helices. When the entire recognition site is only four or six base pairs in length, symmetry-related  $\alpha$ -helices of a protein dimer would collide if each helix had to fit into the major groove. This problem could be avoided if the ends of the helices were used.

This research was funded by NIH grants GM 22526 and GM 29109 to Mark Ptashne. We thank Mark Ptashne, Andrea Jeffrey, Steve Harrison, Don Wiley, Doug Rees, Robert Sauer and Jiahui Wang for support, assistance, and advice. C.O.P. was supported by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research and M.L. was supported by a fellowship from the American Cancer Society.

Received 2 February; accepted 13 May 1982.

1. Ptashne, M. *et al.* *Cell* **19**, 1-11 (1980).
2. Johnson, A. D. *et al.* *Nature* **294**, 217-223 (1981).
3. Sauer, R. T. & Anderregg, R. *Biochemistry* **17**, 1092-1100 (1978).
4. Pabo, C. O., Sauer, R. T., Sturtevant, J. M. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1608-1612 (1979).
5. Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M. & Backman, K. C. *Nature* **279**, 396-400 (1979).
6. Johnson, A. D. thesis, Harvard Univ. (1980).
7. Pabo, C. O. thesis, Harvard Univ. (1980).
8. Bricogne, G. *Acta crystallogr.* **A32**, 832-847 (1976).
9. Diamond, R. *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Venkatesan, K.) (Indian Academy of Science (1981).
10. Maniatis, T. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1531-1535 (1973).

11. Humayun, Z., Kleid, D. & Ptashne, M. *Nucleic Acids Res.* **4**, 1595-1607 (1977).
12. Johnson, A. D. *J. mol. Biol.* (in the press).
13. Pabo, C. O., Krovatin, W., Jeffrey, A. & Sauer, R. T. *Nature* **298**, 441-443 (1982).
14. Sauer, R. T. thesis, Harvard Univ. (1979).
15. Anderson, W. F., Ohlendorf, D. H., Takeda, Y. & Matthews, B. W. *Nature* **290**, 754-758 (1981).
16. Hsiang, M. W., Cole, R. D., Takeda, Y. & Echols, H. *Nature* **270**, 275-277 (1977).
17. Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M., & Pabo, C. O. *Nature* **298**, 447-451 (1982).
18. McKay, D. B. & Steitz, T. A. *Nature* **290**, 744-749 (1981).
19. Zubay, G. & Doty, P. *J. molec. Biol.* **1**, 1-20 (1959).
20. Sung, M. T. & Dixon, G. H. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1616-1623 (1970).
21. Adler, K. *et al.* *Nature* **237**, 322-327 (1972).
22. Warrant, R. W. & Kim, S.-H. *Nature* **271**, 130-135 (1978).
23. Hol, W. G. J., van Duijn, P. T. & Berendsen, H. J. C. *Nature* **273**, 443-446 (1978).

# Homology among DNA-binding proteins suggests use of a conserved super-secondary structure

R. T. Sauer\*, R. R. Yocum†, R. F. Doolittle‡, M. Lewis† & C. O. Pabo†

\* Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

‡ Department of Chemistry, University of California, San Diego, La Jolla, California 92093, USA

*The amino acid sequences of the repressor and cro proteins of phages  $\lambda$ , 434 and P22 are homologous, especially in a region in which repressor and  $\lambda$  cro have a similar  $\alpha$ -helix-turn- $\alpha$ -helix secondary structure. Model-building studies indicate that this structure is important in DNA binding, and we suggest it may be a common feature of many DNA-binding proteins.*

THE related temperate phages  $\lambda$ , 434 and P22<sup>1-3</sup> encode a number of sequence-specific DNA-binding proteins, including a repressor and a cro protein<sup>4,5</sup> which regulate phage transcription. The cro proteins are small, single domain proteins, whereas the repressors contain an N-terminal domain that binds to operator DNA and a C-terminal domain that mediates oligomerization<sup>5-7</sup>. Amino acid sequences of the repressor and cro proteins of  $\lambda$ , 434 and P22 are known (refs 8-15 and R.R.Y., J. Anderson and R.T.S., unpublished), as are the three-dimensional structures of  $\lambda$  cro and of the operator-binding domain of  $\lambda$  repressor<sup>16,17</sup>. Thus these phage proteins provide

an excellent system for studying protein-DNA interactions at the molecular level.

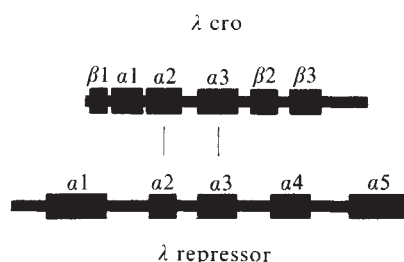
Here, we compare the sequences of the phage repressor and cro proteins in light of the structural information now available. We find a conserved region in these and other DNA-binding proteins and discuss the implications for structure, DNA binding and evolution of these proteins.

## Repressor-cro homologies

Pairwise comparisons of the amino acid sequences of the cro proteins and repressor N-terminal domains from phages  $\lambda$ , 434

| Sequence pair         | Best alignments |      |           | Restricted alignment |      |           |
|-----------------------|-----------------|------|-----------|----------------------|------|-----------|
|                       | No. of IDs      | Gaps | $n\sigma$ | No. of IDs           | % ID | $n\sigma$ |
| 34 cro-434 Rep        | 33              | 0    | 18.6      | 33                   | 45   | 20.3      |
| 34 cro-P22 Rep        | 21              | 1    | 7.2       | 21                   | 30   | 9.2       |
| 34 Rep-P22 Rep        | 24              | 2    | 38.2      | 21                   | 29   | 7.4       |
| 34 Rep- $\lambda$ Rep | 16              | 0    | 3.3       | 18                   | 24   | 5.7       |
| 34 cro- $\lambda$ cro | 15              | 1    | 2.9       | 14                   | 20   | 5.4       |
| 22 Rep-P22 cro        | 18              | 2    | 4.5       | 13                   | 21   | 5.0       |
| 34 Rep- $\lambda$ cro | 11              | 1    | 3.2       | 13                   | 18   | 3.6       |
| Rep-434 cro           | 12              | 0    | 1.6       | 12                   | 16   | 2.7       |
| Rep-P22 cro           | 11              | 0    | 1.9       | 12                   | 20   | 2.3       |
| Rep- $\lambda$ cro    | 12              | 1    | -0.3      | 11                   | 15   | 2.0       |
| 22 Rep- $\lambda$ cro | 9               | 0    | 1.8       | 11                   | 15   | 2.9       |
| 34 cro-P22 cro        | 11              | 0    | 2.2       | 10                   | 16   | 2.1       |
| 22 Rep- $\lambda$ Rep | 17              | 2    | 0.4       | 10                   | 13   | -0.1      |
| cro-P22 cro           | 15              | 2    | 1.7       | 9                    | 15   | 1.1       |
| 34 Rep-P22 cro        | 8               | 0    | -0.3      | 9                    | 14   | 1.9       |





**Fig. 2** Comparison of secondary structures for  $\lambda$  repressor and  $\lambda$  cro. The initial alignment is taken from Fig. 1, but positions of  $\alpha$ -helices and  $\beta$ -sheets are shown<sup>16,17</sup>.

repressor protein has two domains<sup>22-25</sup>; the N-terminal domain binds operator DNA and the C-terminal domain mediates oligomerization. The analogous functions of the Lac repressor domains and the phage repressor domains have been previously cited as evidence of the relatedness of these proteins<sup>26</sup>. Gal repressor shares strong sequence homology with Lac repressor<sup>27</sup>, and presumably also contains two domains.

In Fig. 3, sequences from the lysogeny establishment proteins, and the Lac and Gal repressors have been aligned with the phage repressor and cro protein sequences. The region of comparison shown in Fig. 3 spans the conserved helix-turn-helix region of  $\lambda$  repressor and  $\lambda$  cro. Statistical tests indicate that all of the sequences included in Fig. 3 are significantly related to the parent set.

In Fig. 3, four positions are highly conserved. These correspond to  $\lambda$  repressor residues 33 (7/11 Gln), 37 (10/11 Ala), 41 (11/11 Gly) and 47 (11/11 Val or Ile). Three of these positions are involved in determining the relationship between  $\alpha$ -helices 2 and 3 in  $\lambda$  repressor. Gly 41 forms part of the turn between the helices. We assume that the absence of a side chain at this position is significant and note that glycine is required to form certain types of reverse turns<sup>28</sup>. The Ala 37 and Val 47 side chains make significant van der Waals contacts with each other, and these contacts help determine the angle relating  $\lambda$  repressor helices 2 and 3. Thus, the most highly conserved residues among all of the aligned proteins would serve important structural roles in the helix-turn-helix structures like that found in  $\lambda$  repressor. At other positions, there is a correspondence of residue type with side chain location in  $\lambda$  repressor. Residues 36, 40, 42 and 50, which are generally hydrophobic, form part of the hydrophobic core of  $\lambda$  repressor's N-terminal domain, while residues 33-35, 38-39, 43-46 and 48-49, which are generally hydrophilic, are solvent exposed. These observations suggest that all the repressor, cro and positive activator proteins could readily form similar structures, and supports the idea that the sequence homologies also reflect structural homologies.

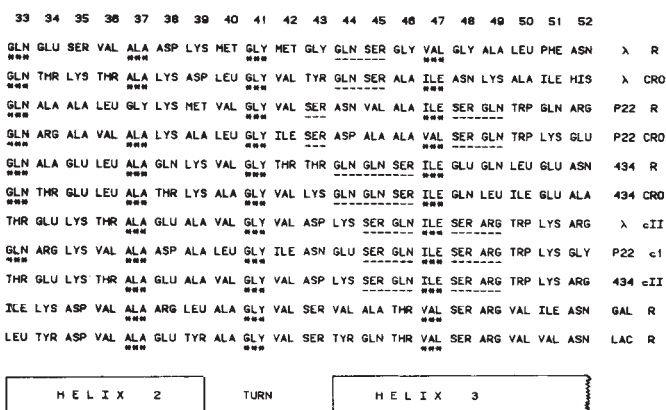
In the case of Lac repressor, biochemical and genetic data support the possibility that residues 6-25 (shown in Fig. 3) form a helix-turn-helix structure, which functions in operator binding. Chemical modification experiments suggest that tyrosines 7, 12 and 17 of Lac repressor are solvent exposed<sup>29,30</sup>, and the corresponding residues of  $\lambda$  repressor (residues 34, 39 and 44) are also solvent exposed<sup>17</sup>. NMR studies further indicate that tyrosines 7 and 17 of Lac repressor are stacked<sup>31</sup>, and we note that the corresponding side chains of  $\lambda$  repressor (Glu 34 and Gln 44) are within hydrogen-bonding distance of each other<sup>17</sup>. At Lac repressor residues 16-19, missense or suppressed nonsense mutations which abolish operator binding have been isolated<sup>32,33</sup>. These residues are therefore excellent candidates for donating side chains to make specific operator contacts. Adler *et al.*<sup>34</sup> originally suggested that these residues formed part of an  $\alpha$ -helical region involved in operator recognition and Beyreuther<sup>24</sup> suggested that residues 16-19 were on the surface of the DNA-binding site. The sequence homologies reported here in combination with the  $\lambda$  repressor and  $\lambda$  cro structures support these ideas.

## C-terminal homologies among the phage repressors

The phage repressor proteins share extensive sequence homology in their C-terminal domains. Comparisons of the C-terminal sequences of the  $\lambda$  and P22 repressors first indicated that the repressors of different phages were related<sup>11</sup>. In Fig. 4, we have extended this comparison to include the 434 and LexA repressors<sup>35</sup>. LexA gene product is a bacterial protein that represses transcription of RecA and other genes implicated in the response to cellular DNA damage<sup>36-38</sup>.

The C-terminal homology among the four repressors (shown in Fig. 4) is highly significant. In pairwise comparisons of these sequences, 434 repressor and P22 repressor are most similar. Their alignment requires no insertions or deletions and 82 of the 139 sequence positions are occupied by identical amino acids. The P22 and  $\lambda$  repressor sequences are identical at 62 of 138 positions, while 434 repressor and  $\lambda$  repressor are identical at 48 of 138 positions. For the three repressors, 44 of the 139 C-terminal positions are identical and many other positions are occupied by chemically similar amino acids. LexA shares 30 identities with 434 repressor, 31 with P22 repressor and 29 with  $\lambda$  repressor. For the four repressors, 20 C-terminal positions are identical. These homologies indicate that LexA repressor is related to the phage repressors.

The phage repressors and LexA are cleaved by activated RecA protein<sup>39-42</sup> in a reaction that may require initial binding of RecA to the repressor C-terminal domain<sup>43</sup>. All four repressors are cleaved at identical Ala-Gly sequences (positions 111-112 of Fig. 4) near the beginning of the C-terminal domain. RecA protein shares sequence homology with LexA and the phage repressors (data not shown), and thus may be a member of this related group of two-domain proteins. The C-terminal domains of the phage repressors also contain the strong repressor oligomerization sites<sup>6</sup>, and are the sites of recognition by the phage P22 antirepressor (J. DeAnda and R.T.S., unpublished).



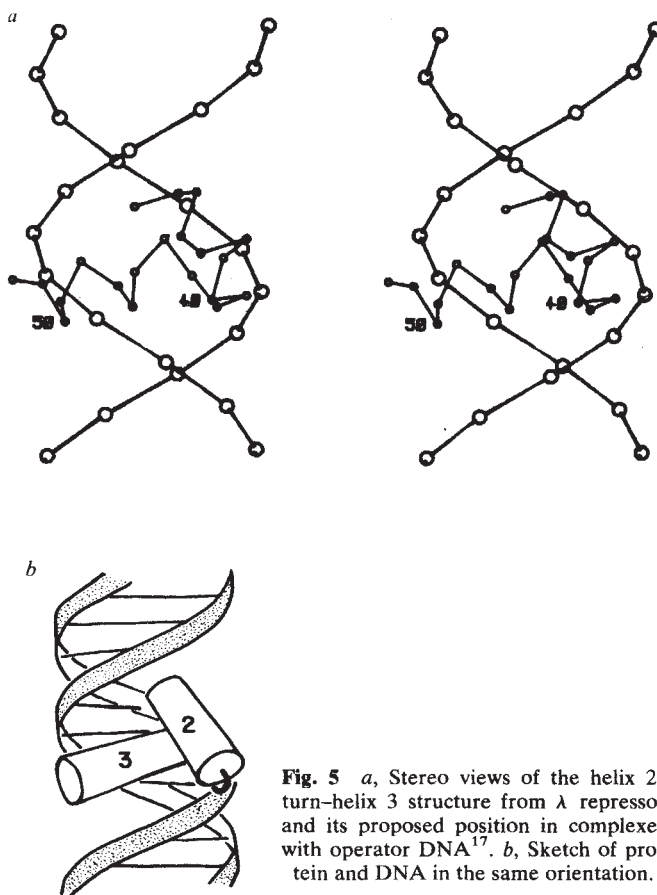
**Fig. 3** Alignment of protein sequences in regions corresponding to  $\alpha$ -helices 2 and 3 in  $\lambda$  repressor and  $\lambda$  cro. Optimal alignments were determined by inspection. Numbering refers to  $\lambda$  repressor. Residues marked by \*\*\* are highly conserved. Underlined residues are predicted to be DNA contact residues in proteins that recognize the same or similar DNA-binding sites (see text). Sequence references not given in legend to Fig. 1: 434 cII, 12;  $\lambda$  cII, 13, 14; P22 cI, R.T.S., unpublished; Lac repressor, 24, 49; Gal repressor, 27. Comparison of any of these sequences with the remaining sequences yields 43-57 identities. The random expectation (corrected for the observed distribution of amino acids in large data bases) is 10.8 identities (95%, range 1-21). Thus all of the sequences are significantly related to the parent set. Homologies (although not necessarily those shown here) between  $\lambda$  cro and 434 cro,  $\lambda$  cII and 434 cII, and Lac, Gal,  $\lambda$  repressor, and P22 repressor have been reported elsewhere<sup>11,12,27</sup>.



## Evolution of phage repressors and *cro*

The sequence homologies noted in the previous sections clearly show that the repressor and *cro* proteins of phages  $\lambda$ , 434 and P22 have evolved from a common ancestor. The repressors may have evolved from an ancestral *cro* or single-domain DNA-binding protein by gene duplication or by recombinational addition of a C-terminal domain. Alternatively, the one-domain *cro* proteins may have evolved from a repressor by loss of its C-terminal domain. The sequence homologies do not distinguish between these two possibilities, but do suggest that current *cro* and repressor proteins are not linear descendants of an event that created an ancestral *cro* and ancestral repressor. In this case, the *cro* proteins should be most similar to other *cro*s, and the repressors should be most similar to other repressors. Instead, the *cro* proteins of 434 and P22 are more related to their own repressors than to the other *cro* proteins (Table 1), suggesting that several independent instances of cross evolving from repressors or vice versa must have occurred.

The repressor N-terminal domains, the *cro* proteins and the activator proteins all bind to specific DNA sequences, and thus from a functional point of view their common origin is not surprising. In the cases where it is known, all of these proteins bind to a DNA site which spans approximately two turns of the double helix. In fact, *cro* and repressor from any one phage bind to the same operator sites<sup>44,45</sup>. The N-terminal repressor domains and the activator proteins also share the common function of positively regulating transcription<sup>7,21</sup>. A mutant  $\lambda$  repressor which binds operator DNA but fails to regulate trans-



**Fig. 5** *a*, Stereo views of the helix 2-turn-helix 3 structure from  $\lambda$  repressor and its proposed position in complexes with operator DNA<sup>17</sup>. *b*, Sketch of protein and DNA in the same orientation.

|                             |                                                         |           |     |           |   |
|-----------------------------|---------------------------------------------------------|-----------|-----|-----------|---|
| 95                          | 100                                                     | 105       | 110 |           |   |
| MET GLN PRO SER LEU ARG SER | GLU TYR GLU TYR PRO VAL PHE SER HIS VAL GLN ALA GLY     | $\lambda$ | R   |           |   |
| HIS SER ARG HIS GLU PRO     | ARG GLY SER TYR PRO LEU ILE SER TRP VAL SER ALA GLY     | P22       | R   |           |   |
| VAL GLY HIS VAL GLU PRO     | LYS GLY LYS TYR PRO LEU ILE SER MET VAL ARG ALA GLY     | 434       | R   |           |   |
| ARG LEU LEU GLN GLU GLU     | GLU GLY LEU PRO LEU VAL GLY ARG VAL ALA ALA GLY         | LEX       | R   |           |   |
| 115                         | 120                                                     | 125       | 130 |           |   |
| MET PHE SER PRO GLU LEU ARG | THR PHE THR LYS GLY ASP ALA GLU ARG TRP VAL SER THR     | $\lambda$ | R   |           |   |
| GLN TRP MET GLU ALA VAL     | GLU PRO TYR HIS LYS ARG ALA ILE GLU ASN TRP HIS ASP THR | P22       | R   |           |   |
| SER TRP CYS GLU ALA CYS     | GLU PRO TYR ASP ILE LYS ASP ILE ALA GLU TRP TYR ASP SER | 434       | R   |           |   |
| GLU PRO LEU LEU ALA GLN     | GLN HIS ILE GLU GLY HIS TYR GLN VAL ASP PRO SER LEU PHE | LEX       | R   |           |   |
| 135                         | 140                                                     | 145       | 150 |           |   |
| THR LYS LYS ALA SER ASP     | SER ALA PHE TRP LEU GLU VAL GLU GLY ASN SER MET THR ALA | $\lambda$ | R   |           |   |
| THR VAL ASP CYS SER GLU     | ASP SER PHE TRP LEU ASP VAL GLN GLY ASP SER MET THR ALA | P22       | R   |           |   |
| ASP VAL ASN LEU LEU GLY     | ASN GLY PHE TRP LEU LYS VAL GLU GLY ASP SER MET THR SER | 434       | R   |           |   |
| LYS PRO ASN ALA ASP         | PHE LEU LEU ARG VAL SER GLY MET SER MET LYS ASP         | LEX       | R   |           |   |
| 155                         | 160                                                     | 165       | 170 |           |   |
| PRO THR GLY SER LYS PRO     | SER PHE PRO ASP GLY MET LEU ILE LEU VAL ASP PRO GLU GLN | $\lambda$ | R   |           |   |
| PRO ALA GLY LEU             | SER ILE PRO GLU GLY MET ILE ILE LEU VAL ASP PRO GLU VAL | P22       | R   |           |   |
| PRO VAL GLY GLN             | SER ILE PRO GLU GLY HIS MET VAL LEU VAL ASP THR GLY ARG | 434       | R   |           |   |
| ILE GLY                     | ILE MET ASP GLY ASP LEU LEU ALA VAL HIS LYS THR GLN     | LEX       | R   |           |   |
| 175                         | 180                                                     | 185       | 190 |           |   |
| ALA VAL GLU PRO GLY ASP     | PHE CYS ILE ALA ARG LEU GLY GLY ASP GLU PHE THR PHE     | $\lambda$ | R   |           |   |
| GLU PRO ARG ASN GLY LYS     | LEU VAL VAL ALA LYS LEU GLU GLY GLU ASN GLU ALA THR PHE | P22       | R   |           |   |
| GLU PRO VAL ASN GLY SER     | LEU VAL VAL ALA LYS LEU THR ASP ALA ASN GLU ALA THR PHE | 434       | R   |           |   |
| ASP VAL ARG ASN GLY GLN     | VAL VAL VAL ALA ARG ILE ASP ASP GLU VAL THR VAL         | LEX       | R   |           |   |
| 195                         | 200                                                     | 205       | 210 |           |   |
| LYS LYS LEU ILE ARG ASP     | SER GLY GLN VAL PHE LEU GLN PRO LEU ASN PRO GLN TYR PRO | $\lambda$ | R   |           |   |
| LYS LYS LEU VAL MET ASP     | ALA GLY ARG LYS PHE LEU LYS PRO LEU ASN PRO GLN TYR PRO | P22       | R   |           |   |
| LYS LYS LEU VAL ILE ASP     | GLY GLY GLN LYS TYR LEU LYS GLY LEU ASN PRO SER TRP PRO | 434       | R   |           |   |
| LYS ARG LEU LYS LYS GLN     | GLY ASN LYS VAL GLU LEU LEU PRO GLU ASN SER GLU PHE LYS | LEX       | R   |           |   |
| 215                         | 220                                                     | 225       | 230 |           |   |
| MET ILE PRO CYS ASN GLU     | SER CYS SER VAL VAL GLY LYS VAL ILE ALA SER GLN TRP PRO | $\lambda$ | R   |           |   |
| MET ILE GLU ILE ASN GLY     | ASN CYS LYS ILE ILE GLY VAL VAL VAL ASP ALA LYS LEU ALA | P22       | R   |           |   |
| MET THR PRO ILE ASN GLY     | ASN CYS LYS ILE ILE GLY VAL VAL VAL ASP ALA ARG VAL LYS | 434       | R   |           |   |
| PRO ILE VAL VAL ASP LEU     | ARG GLN GLN SER PHE THR ILE GLU GLY LEU ALA VAL GLY VAL | LEX       | R   |           |   |
| 235                         |                                                         |           |     |           |   |
| GLU GLU THR PHE GLY COOH    |                                                         |           |     | $\lambda$ | R |
| ASN LEU PRO COOH            |                                                         |           |     | P22       | R |
| PHE VAL COOH                |                                                         |           |     | 434       | R |
| ILE ARG ASN GLY ASP TRP     | LEU COOH                                                |           |     | LEX       | R |

**Fig. 4** Alignment of repressor C-terminal domain sequences. Numbering is that of  $\lambda$  repressor. Sequence references given in legends to Figs 2 and 3. Residue identities are indicated by underlining.

cription positively has been isolated<sup>46</sup>. This mutation changes residue 43 of  $\lambda$  repressor and thus occurs in the region of the proposed helix-turn-helix structure where the activator and repressor proteins are most similar.

## Discussion

Figure 5 depicts the helix-turn-helix structure from  $\lambda$  repressor and shows its suggested relationship to operator DNA<sup>17</sup>. The helix-turn-helix structure found in  $\lambda$  *cro* assumes a similar but slightly different orientation with respect to the operator in the suggested complex<sup>16</sup>. Figure 5 shows that residues at the N-terminal part of helix 3 of  $\lambda$  repressor could contact sites in the major groove, and shows that the N-terminal end of helix 2 is also in a position to contact the DNA. In  $\lambda$  *cro*, the analogous positions at the N-terminal ends of helices 2 and 3 are also close to the operator, but in addition, residues from the middle of helix 3 are in a position to make major groove contacts. Clearly, in any particular case, the orientation of the  $\alpha$ -helices with respect to the DNA will determine which side chains can contact the operator. However, the  $\lambda$  repressor and  $\lambda$  *cro* examples suggest that residues corresponding to positions 33–34, 43–46 and 48–49 in Fig. 3 will be potential DNA contact residues in many of the aligned proteins. As noted above, residue side chains at these positions are generally hydrophilic and should therefore be capable of forming hydrogen bonds with the DNA.

Because the repressor and *cro* from any one phage bind to the same operator sequences in the phage DNA, it is reasonable to assume that some of the specific operator contacts made by this repressor and *cro* will be made by homologous amino acid residues. For the phages discussed, we predict that the sequences Gln-Ser ( $\lambda$ ), Gln-Gln-Ser (434) and Ser-(x)<sub>4</sub>-Ser-Gln (P22) in helix 3 could participate in such operator contacts. In other cases, where the DNA recognition sites of two proteins differ, homologies in 'contact residues' may indicate that side chains at these positions hydrogen bond to the same or similar bases. Thus, the fact that  $\lambda$ c II, 434c II and P22c 1 gene products

have identical Ser-Gln-x-Ser-Arg sequences at analogous positions in helix 3 suggests to us that their DNA recognition sites will show some homology.

The sequence alignment of the phage repressor and cro proteins (Fig. 1) allows us to draw another conclusion concerning the mechanisms by which these proteins recognize specific operator DNA sequences. The  $\lambda$  repressor contains a long N-terminal extension when compared with the other proteins. The N-terminal residues of this extension form a flexible arm which wraps around the operator to contact bases on the backside of the operator when repressor binds (see accompanying article<sup>47</sup>). The  $\lambda$  cro protein does not make these contacts on the back of an operator site<sup>44,45</sup>, nor does it have an analogous arm<sup>16</sup>. As repressor's 'arm' is formed by residues with no analogue in the other cro and repressor proteins, we predict that these other proteins will not contain N-terminal 'arms' that function in operator DNA binding.

Received 2 February; accepted 13 May 1982.

- Hershey, A. D. & Dove, W. in *The Bacteriophage  $\lambda$*  (ed. Hershey, A. D.) 3-12 (Cold Spring Harbor Laboratory, New York, 1971).
- Botstein, D. & Herskowitz, I. *Nature* **251**, 584-589 (1974).
- Botstein, D. *Ann. N.Y. Acad. Sci.* **354**, 484-491 (1980).
- Ptashne, M. *et al. Cell* **19**, 1-11 (1980).
- Johnson, A. D. *et al. Nature* **294**, 217-223 (1981).
- Pabo, C. O., Sauer, R. T., Sturtevant, J. M., & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1608-1612 (1979).
- Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M. & Backman, K. C. *Nature* **279**, 396-400 (1979).
- Sauer, R. T. & Anderregg, R. *Biochemistry* **17**, 1092-1100 (1978).
- Roberts, T. M., Shimatake, H., Brady, C. & Rosenberg, M. *Nature* **270**, 274-275 (1977).
- Hsiang, M. W., Cole, R. D., Takeda, Y. & Echols, H. *Nature* **270**, 275-277 (1977).
- Sauer, R. T. *et al. Biochemistry* **20**, 3591-3598 (1981).
- Grosschedl, R. & Schwartz, E. *Nucleic Acids Res.* **6**, 867-881 (1979).
- Schwarz, E., Scherer, G., Hobom, G. & Kossel, H. *Nature* **272**, 410-414 (1978).
- Ovchinnikov, Y. A. *et al. Gene* **6**, 235-249 (1979).
- Sauer, R. T. *Nature* **276**, 301-302 (1978).
- Anderson, W. F., Ohlendorf, D. H., Takeda, Y. & Matthews, B. W. *Nature* **290**, 754-758 (1981).
- Pabo, C. O. & Lewis, M. *Nature* **298**, 443-447 (1982).
- Needleman, S. B. & Wunch, C. D. *J. molec. Biol.* **48**, 443-453 (1970).
- Doolittle, R. F. *Science* **214**, 149-159 (1981).
- Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. in *The Operon* (eds Miller, J. H. & Reznikoff, W. S.) 345-371 (Cold Spring Harbor Laboratory, New York, 1978).
- Shimatake, H. & Rosenberg, M. *Nature* **292**, 128-132 (1981).
- Geisler, N. & Weber, K. *Biochemistry* **16**, 938-943 (1977).
- Ogata, R. T. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5851-5854 (1978).
- Beyreuther, K. in *The Operon* (eds Miller, J. H. & Reznikoff, W. S.) 123-154 (Cold Spring Harbor Laboratory, New York, 1978).
- Weber, K. & Geisler, N. in *The Operon* (eds Miller, J. H. & Reznikoff, W. S.) 155-175 (Cold Spring Harbor Laboratory, New York, 1978).

We thank Brian Matthews, Wayne Anderson, Benno Muller-Hill and Tom Steitz for communication of results before publication. Matthews and his colleagues have independently suggested that a conserved  $\alpha$ -helical secondary structure is used by many DNA-binding proteins<sup>48</sup>, and a comparison of  $\alpha$ -carbon positions has shown that helices 2 and 3 of  $\lambda$  cro are related to helices E and F of CAP<sup>50</sup>. We thank Mark Ptashne for encouragement, Patty Rich for help in preparation of this manuscript, Roger Brent for drawing our attention to the lex A repressor-phage repressor homology, and Wayne Anderson for drawing our attention to the  $\lambda$  cro- $\lambda$  cII homology. This research was supported by NIH grants to R.T.S. (AI-16892), Mark Ptashne (GM-22526), and R.F. Doolittle (HL-26873). R.R.Y. was a postdoctoral fellow of the Helen Hay Whitney Foundation. C.O.P. was a postdoctoral fellow of the Jane Coffin Childs Memorial Fund for Medical Research. M.L. was a postdoctoral fellow of the American Cancer Society.

- Muller-Hill, B., Gronenborn, B., Kania, J., Schlotmann, M. & Beyreuther, K. in *Nucleic Acid-Protein Recognition* (ed. Vogel, H. J.) 219-236 (Academic, New York, 1977).
- Wilcken-Bergmann, B. V. & Muller-Hill, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2427-2431 (1981).
- Schultz, G. E. & Schirmer, R. H. *Principles of Protein Structure* (Springer, New York, 1979).
- Fanning, T. G. *Biochemistry* **14**, 2512-2520 (1975).
- Alexander, M. E., Burgum, A. A., Noall, R. A., Shaw, M. D. & Matthews, K. S. *Biochim. biophys. Acta* **493**, 367-379 (1977).
- Arndt, K. T., Boschelli, F., Lu, P., & Miller, J. H. *Biochemistry* **20**, 6109-6118 (1981).
- Miller, J. H. in *The Operon* (eds Miller, J. H. & Reznikoff, W. S.) 31-88 (Cold Spring Harbor Laboratory, New York, 1978).
- Weber, K., Platt, T., Ganem, D. & Miller, J. H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3624-3628 (1972).
- Adler, K. *et al. Nature* **237**, 322-327 (1972).
- Horii, T., Ogawa, T. & Ogawa, H. *Cell* **23**, 689-698 (1981).
- Mount, D. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 300-304 (1977).
- Kenyon, C. J. & Walker, G. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2813-2819 (1980).
- Brent, R. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4204-4208 (1981).
- Horii, T. *et al. Cell* **27**, 515-522 (1981).
- Roberts, J. W., Roberts, C. W. & Craig, N. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4714-4718 (1978).
- Phizicky, E. & Roberts, J. J. *molec. Biol.* **139**, 319-328 (1980).
- Little, J. W., Edmiston, S. H., Pacelli, L. Z. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3225-3229 (1980).
- Sauer, R. T., Ross, M. J. & Ptashne, M. *J. biol. Chem.* **257**, 4458-4462 (1982).
- Johnson, A., Meyer, B. J., Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1783-1787 (1978).
- Johnson, A. thesis, Harvard Univ. (1980).
- Guarente, L., Nye, J. S., Hochschild, A. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2236-2239 (1982).
- Pabo, C. O., Krovatin, W., Jeffrey, A. & Sauer, R. T. *Nature* **298**, 441-443 (1982).
- Matthews, B. W., Ohlendorf, D. H., Anderson, W. F. & Takeda, Y. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1428-1432 (1982).
- Beyreuther, K. *et al. Eur. J. Biochem.* **59**, 491-509 (1975).
- Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3097-3100 (1982).

## LETTERS TO NATURE

### Is the Universe rotating?

P. Birch

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

From the study of the position angles and polarization of high luminosity classical-double radio sources, it appears that the difference between the position angles of elongation and of polarization are highly organized, being generally positive in one half of the sky and negative in the other. The effect was first noticed amongst a sample of 94 3CR sources and later confirmed in three independent samples. Such a phenomenon can only have a physical explanation on a cosmic scale; an attractive theory is that it demonstrates the existence of a universal vorticity, that is, that the Universe is rotating with an angular velocity  $\sim 10^{-13}$  rad yr<sup>-1</sup>. This would have drastic cosmological consequences, since it would violate Mach's principle<sup>1,2</sup> and the widely held assumption of large-scale isotropy.

A complete sample of 94 3CR radio sources in the declination range  $15^\circ < \delta < 40^\circ$  has been observed at Jodrell Bank at four

wavelengths; detailed polarization results are to be published elsewhere. These results include values of  $\Delta$ , which is the position angle of the magnetic field vectors minus the position angle of elongation of the source (the major axis); obtained for 45 of these sources (Table 1a). The sources, ordered in right ascension (RA), fell into two groups divided at RA 0915 and 2115, in which the sign of  $\Delta$  became, in sequence:

Sample 1: -----+-----+-----+ 4(+), 18(-)

Sample 2: ++++++-----+-----+ 17(+), 6(-)

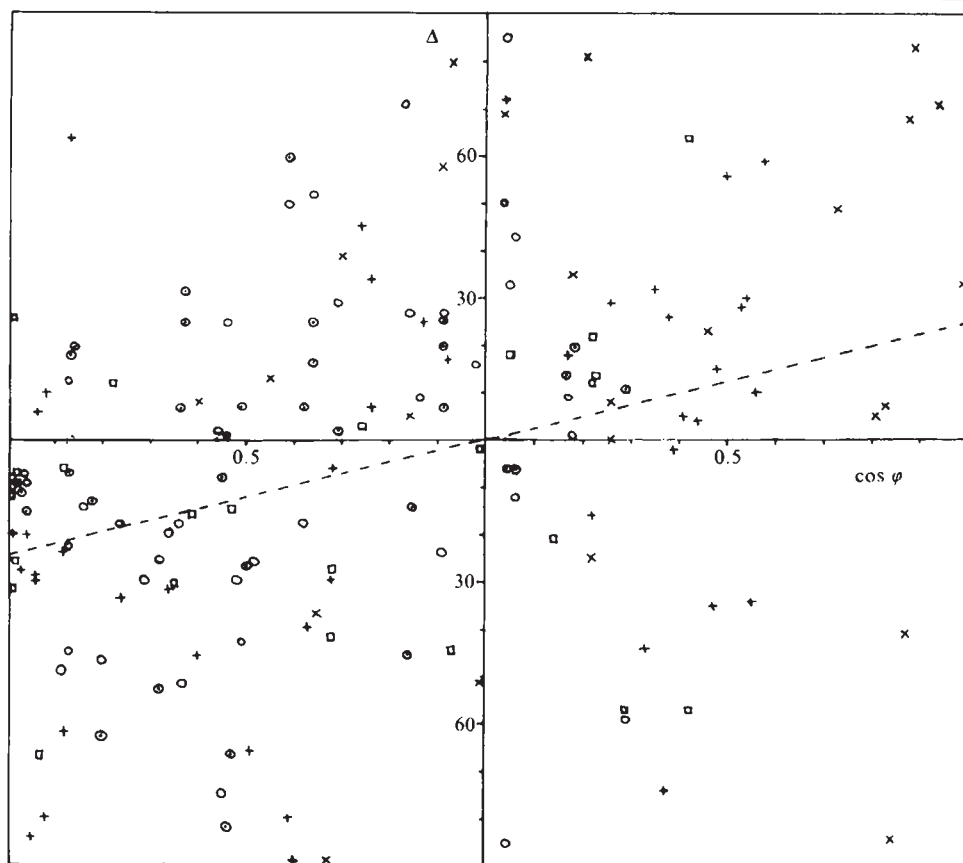
The difference between the two was striking, and prompted further analysis of the data; even taking account of the arbitrary division in RA, this result was already significant, the  $\chi^2$  test for two degrees of freedom yielding a significance level,  $\alpha \approx 2.6 \times 10^{-3}$  (that is, the probability that the difference between the samples was due to chance was  $\leq 1\%$ ).

If the magnitude as well as the sign of  $\Delta$  is considered the following mean values are obtained:

Sample 1:  $\bar{\Delta}_1 \approx -31.6^\circ \pm 8.0^\circ$

Sample 2:  $\bar{\Delta}_2 \approx +12.3^\circ \pm 7.3^\circ$

These differ at the significance level,  $\alpha \approx 2.5 \times 10^{-4}$ . The quoted



**Fig. 1** A plot of  $\Delta$  against  $\cos \phi$ :  $\Delta$  is defined as the position angle of the magnetic field vector minus the position angle of the major axis of the radio source;  $\phi$  is defined as the angular separation on the sky of the radio source and the pole of the dipole anisotropy in  $\Delta$ , which has been obtained by a least-squares fit to the  $\Delta$  values plotted. Here, the pole position is RA 14 h 55 min, dec  $-35^\circ$ ; and the expected magnitude of  $\Delta$  is given by the dotted line,  $\langle \Delta \rangle = 25^\circ \cos \phi$ . Also plotted are values of  $\Delta_i$ , which is the position angle of elongation of the components minus the position angle of the major axis of the source. The symbols for the four samples are: PB sample, +; Laing<sup>6</sup> sample, ○ and (for  $\Delta_i$ ) ○; Ekers<sup>7</sup> sample, ×; and Conway<sup>8</sup> sample, □.

uncertainties are  $\Delta_{r.m.s.} (N-1)^{-1/2}$  where  $\Delta_{r.m.s.}$  ( $\approx 35^\circ$ ) is the root mean square deviation from the mean of the values of  $\Delta$  and  $N$  is the number of sources in each sample. The spread in  $\Delta$  is mainly intrinsic, only  $\sim 10^\circ$  arising from uncertainty in the individual values.

The values of  $\Delta$  in Table 1a were derived by averaging the values obtained for the outer components of each source; in general, these agree with the values derived from the integrated polarization alone, for which the effect is present at a similar significance level,  $\alpha \approx 4.6 \times 10^{-4}$ . The polarization data, and the values obtained for the intrinsic position angle of the magnetic field, agree well with previously published measurements<sup>3,4</sup>, and according to accepted criteria<sup>3,5</sup> there was no ambiguity in the rotation measure fits for the 45 sources for which values of  $\Delta$  were obtained, all ambiguous cases (there were 10) having been rejected.

The results of the division in RA suggest the presence of a dipole anisotropy in  $\Delta$ . If the angular separation of the radio source and the pole of anisotropy is  $\phi$ , then the magnitude of the dipole term is proportional to  $\cos \phi$ . A least-squares fit to a dipolar anisotropy yields a pole in the direction  $\sim$ RA 1320, dec  $-40^\circ$ , and a magnitude  $\langle \Delta \rangle = (35^\circ \pm 6^\circ) \cos \phi$ .

The effect has been sought in other samples<sup>6-8</sup> which together cover most of the sky with the exception of the galactic plane. It does not appear that the manner in which the sources were originally selected for the samples can have introduced any position-dependent bias of position angles. The values of  $\Delta$  for each of the three new independent samples, derived using published polarization measurements<sup>3</sup>, are listed in Table 1. When divided at RA 0915 and 2115, and tested against the dipole model given above, they confirm it at high levels of significance. Not only do the values of  $\bar{\Delta}$  differ from zero and each other, they do so in the way predicted by the model; this is particularly striking in the southern sample, for which the RA dependence is markedly different from the northern sample.

Thus these other samples clearly confirm the existence of a dipole anisotropy in  $\Delta$ ; the probability of this result arising by chance is very small.

We have searched for systematic errors that could explain

this effect. A systematic calibration error could make  $\bar{\Delta} \neq 0$ , but could not reproduce the variation with position on the sky. The effect is about equal to the correction for ionospheric and galactic Faraday rotations and much larger than their uncertainty of a few degrees; since there was no ambiguity in the magnitude of the rotation measures, it is hard to see how any residual errors could be systematic and RA dependent. The direction of the pole of anisotropy is near ( $l$  320°,  $b$  +20°) in galactic coordinates; this suggests that the effect is not of galactic origin, as this direction does not correspond to the direction of the galactic magnetic field<sup>9</sup>, nor to the axis of rotation ( $b = 90^\circ$ ), nor to any other obvious physical axis.

So far the analysis has been concerned with the magnetic field configuration in radio sources. It is, however, known<sup>10</sup> that in high luminosity radio sources the magnetic field orientation tends to follow the contours of brightness (an effect explained by anisotropic compression or shear<sup>11</sup>). The difference between the position angles of elongation of individual components (the direction of the 'tails') and of the major axis of the source  $\Delta_i$ , might therefore be expected to show the same anisotropy as is found in the values of  $\Delta$  (if this is caused by shearing or compression of the components). Maps of sufficient sensitivity and resolution to obtain accurate values of  $\Delta_i$  (for a considerable number of sources) were only available for the Laing sample (out of the four samples in Table 1); using the published data<sup>6</sup>, values of  $\Delta_i$  were derived for 30 of Laing's sources (Table 2). For these sources, divided at RA 09 h 15 min and 21 h 15 min, the model predicts  $\Delta_1 = -17.4^\circ$ ,  $\Delta_2 = +1.7^\circ$  and  $\Delta_i = \Delta$ , whereas the mean values of  $\Delta_i$  obtained are:

$$\text{Sample 1: } \bar{\Delta}_{i1} = -16.9^\circ \pm 8.1^\circ$$

$$\text{Sample 2: } \bar{\Delta}_{i2} = +8.9^\circ \pm 5.7^\circ$$

These values agree with the prediction, and differ from each other as predicted with  $\alpha \approx 4.0 \times 10^{-3}$ . Thus the same effect appears to be present at a similar level of significance. This is important because the values of  $\Delta_i$  do not depend on polarization measurements and are subject to entirely different kinds of error; this rules out explanations based on errors in deriving the magnetic position angles.



The data in Tables 1 and 2 can be brought together into a single sample, which covers most of the sky ( $\sim 80\%$ ). It is well-sampled from a redshift  $z \approx 0.03$  to  $\approx 1.5$ , and when the data are split into high and low  $z$  samples the effect is present in both; no correlation with redshift is observed. The effect is also present equally for radio galaxies and for quasars. All 132 sources were used to refine the estimate of dipole anisotropy by a least-squares fit as above; the overall mean pseudo-vector offset is found to be:

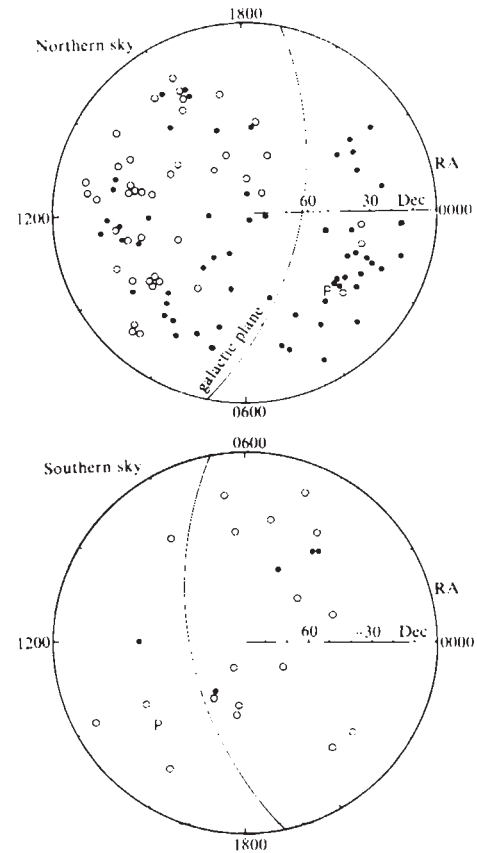
$$\langle \Delta \rangle = (25^\circ \pm 5^\circ), \text{RA } (14 \text{ h } 55 \text{ min} \pm 30 \text{ min}), \text{dec } (-35^\circ \pm 5^\circ)$$

The individual values of  $\Delta$  and  $\Delta_i$  are plotted against  $\cos \phi$  in Fig. 1: the effect is clearly visible as an excess of points in the first and third quadrants. For the same sources Fig. 2 shows the positions on the sky and the sign of  $\Delta$ ; the dipole anisotropy is evident as an excess of negative values around the pole in the northern sky and of positive values around the opposite pole. Standard statistical tests of correlation have been applied to Fig. 1, taking due account of the number of degrees of freedom available in choosing the dipole axis; they yield a significance level,  $\alpha \approx 10^{-8}$ .

An explanation for this phenomenon must involve anisotropy on a cosmic scale (over 10 Gpc or more), yielding a preferred orientation of widely separated radio sources. The phenomenon consists of a handedness in the (magnetic) outline, such that sources tend to look S-shaped in one half of the sky and like the mirror image of an S in the other. This can only be explained by a pseudo-vector rather than a true vector or symmetric tensor effect. Possible causes of the pseudo-vector phenomenon

**Table 1** Magnetic field offset angles for radio sources

| Source                                                  | $\Delta$ | Source   | $\Delta$ | Source   | $\Delta$ | Source  | $\Delta$ |
|---------------------------------------------------------|----------|----------|----------|----------|----------|---------|----------|
| <b>a PB Sample from 3CR</b>                             |          |          |          |          |          |         |          |
| 0017+15                                                 | -34      | 0459+25  | -62      | 1158+31  | -16      | 1618+17 | +10      |
| 0031+39                                                 | +64      | 0659+25  | -46      | 1232+21  | +4       | 1622+23 | +15      |
| 0107+31                                                 | +10      | 0802+24  | -89      | 1241+16  | +56      | 1626+39 | +5       |
| 0123+32                                                 | -29      | 0824+29  | -40      | 1251+15  | +28      | 1627+23 | -35      |
| 0125+28                                                 | -30      | 0903+16  | +25      | 1251+27  | -74      | 1726+31 | +29      |
| 0132+37                                                 | -20      | 0905+38  | -6       | 1308+27  | -2       | 2141+27 | -80      |
| 0133+20                                                 | -80      | 0927+36  | +45      | 1350+31  | +26      | 2145+15 | -30      |
| 0154+28                                                 | -28      | 0936+36  | +7       | 1420+19  | +59      | 2203+29 | -66      |
| 0220+39                                                 | -20      | 0938+39  | +34      | 1529+35  | -44      | 2318+23 | -32      |
| 0300+16                                                 | -84      | 0958+29  | +17      | 1545+21  | +30      |         |          |
| 0307+16                                                 | +6       | 1108+35  | +72      | 1553+20  | -34      |         |          |
| 0453+22                                                 | -24      | 1142+31  | +18      | 1615+32  | +32      |         |          |
| <b>b Laing sample<sup>6</sup> from 3CR</b>              |          |          |          |          |          |         |          |
| 0013+79                                                 | -26      | 0605+48  | -47      | 1140+22  | -59      | 1704+61 | +71      |
| 0017+15                                                 | -18      | 0651+54  | -30      | 1142+31  | +9       | 1825+74 | +50      |
| 0040+51                                                 | -45      | 0733+70  | -18      | 1157+73  | -18      | 1832+47 | -24      |
| 0106+13                                                 | -14      | 0809+48  | +25      | 1206+44  | -85      | 1842+45 | +27      |
| 0154+28                                                 | -7       | 0833+65  | -75      | 1254+47  | +43      | 1845+79 | -26      |
| 0210+86                                                 | -52      | 0835+58  | -30      | 1258+40  | -1       | 1939+60 | +52      |
| 0229+34                                                 | -8       | 0850+14  | +27      | 1409+52  | +85      | 2104+76 | +2       |
| 0410+11                                                 | +13      | 0855+14  | +9       | 1441+52  | +33      | 2153+37 | -43      |
| 0415+37                                                 | -9       | 1030+58  | +29      | 1609+66  | +50      | 2318+23 | -20      |
| 0459+25                                                 | -49      | 1111+40  | +16      | 1658+47  | -12      |         |          |
| <b>c Ekers sample<sup>7</sup> from the southern sky</b> |          |          |          |          |          |         |          |
| 0043-42                                                 | 90       | 0431-133 | +8       | 1358-11  | +83      | 1737-60 | +5       |
| 0114-47                                                 | +58      | 0511-30  | +39      | 1413-36  | +33      | 2040-26 | 90       |
| 0242-57                                                 | +69      | 0518-45  | 90       | 1556-21  | +71      | 2058-28 | 0        |
| 0336-35                                                 | -89      | 0618-37  | +80      | 1602-63  | -41      | 2104-25 | +81      |
| 0334-34                                                 | -37      | 0634-20  | +8       | 1610-608 | +68      | 2153-69 | +23      |
| 0349-27                                                 | +13      | 0819-30  | +35      | 1637-77  | +49      | 2317-27 | +5       |
| 0427-53                                                 | -51      | 1211-41  | -84      | 1733-56  | +7       | 2356-61 | -25      |
| <b>d Conway sample<sup>8</sup> from 4C</b>              |          |          |          |          |          |         |          |
| 0035+38                                                 | -6       | 0349+26  | -9       | 1029+25  | +18      | 1333+27 | +64      |
| 0128+25                                                 | -68      | 0658+35  | -31      | 1100+37  | -2       | 1634+26 | -57      |
| 0211+34                                                 | -32      | 0733+29  | -15      | 1123+30  | -21      | 2239+33 | -16      |
| 0214+27                                                 | -26      | 0840+29  | -42      | 1158+25  | -58      | 2356+27 | +12      |
| 0229+35                                                 | -11      | 0854+34  | -28      | 1253+37  | +12      |         |          |
| 0241+29                                                 | +26      | 0931+39  | +3       | 1257+38  | +13      |         |          |
| 0313+34                                                 | -7       | 0950+25  | -45      | 1301+38  | +22      |         |          |



**Fig. 2** This diagram shows the source positions in right ascension and declination. Sources with negative  $\Delta$  are plotted as  $\bullet$  whereas those with positive  $\Delta$  are plotted as  $\circ$ .  $\Delta$  is defined as the position angle of the magnetic field vector minus the position angle of the major axis of the radio source. The positions of the poles of anisotropy are marked P.

appear to be limited to (1) large-scale magnetic fields, or (2) large-scale or universal rotation.

Measurements of the Faraday rotation of extragalactic radio sources have been used to place limits on the magnitude of a universal magnetic field<sup>12</sup>, which suggests that it would be far too weak to affect the sources' morphology directly. Furthermore, such a field could only produce a pseudovector effect through the inherent asymmetry between electrons and protons, so that strong electric fields and charge separation would be required. A viable explanation along these lines appears unlikely.

An alternative and attractive explanation is that the radio sources rotate relative to the intergalactic medium, the axis of rotation being preferentially aligned with a universal vorticity; for, if the Universe has net angular momentum, this will be conserved and provide a natural axis of rotation aligned over cosmic scales. Considering representative models in which slowly rotating radio galaxies, preferentially aligned, are assumed to have condensed from the intergalactic medium under constant angular momentum, the appropriate angular velocity of the Universe is found to be of the order of  $10^{-13}$  rad yr<sup>-1</sup>. This value is obtained by considering the overall angular momentum of the radio galaxies and the intergalactic medium, and the mean density of the Universe, to determine the mean vorticity.

A vorticity on a smaller scale, in the range of scale size  $1 \leq z \leq 1,000$ , might provide an alternative explanation for the alignments. It would, however, impose radial velocities of order  $300 \text{ km s}^{-1}$  on the microwave background, which would have been detected<sup>13,14</sup>. Note that, in the past, such large-scale vorticity has been somewhat loosely called 'rotation of the Universe', and stringent limits have been placed on the maximum allowable vorticity<sup>13-16</sup>. By contrast, a pure universal rotation

**Table 2** Offset angle of elongation of outer components

| Source $\Delta_1$                  | Source $\Delta_2$ | Source $\Delta_3$ | Source $\Delta_4$ |
|------------------------------------|-------------------|-------------------|-------------------|
| Laing sample <sup>6</sup> from 3CR |                   |                   |                   |
| 0013+79                            | -53 0459+25       | +20 1140+22       | +11 1832+47       |
| 0040+51                            | +18 0605+48       | -63 1142+31       | +14 1842+45       |
| 0106+13                            | -13 0733+70       | +7 1157+73        | +7 1845+79        |
| 0154+28                            | -8 0809+48        | -82 1258+40       | +20 1939+60       |
| 0210+86                            | +31 0833+65       | -8 1609+66        | -6 2104+76        |
| 0229+34                            | -8 0835+58        | -67 1658+47       | -6 2153+37        |
| 0410+11                            | -8 0850+14        | -46 1704+61       | -14               |
| 0415+37                            | -15 1030+58       | +1 1825+74        | +60               |

For 22 of these sources, component position angles were obtained from Laing's Table 7 (ref. 6); for the remainder the position angles were taken from the maps. The effect is still present if only Laing's tabulated position angles are used ( $\alpha \approx 0.037$ )

would produce a pure quadrupole anisotropy<sup>15</sup> well below observational limits<sup>15,16</sup>.

Thus there appears to be strong evidence that the Universe is anisotropic on a large-scale, producing position angle offsets in the polarization and brightness distributions of radio sources. These can probably be explained by a rotation of the Universe, with an angular velocity pseudo-vector:

$$\omega_u = (\sim 10^{-13} \text{ rad yr}^{-1}), \text{ RA } (2 \text{ h } 55 \text{ min } \pm 30 \text{ min}), \text{ dec } (35^\circ \pm 5^\circ)$$

I thank Dr R. G. Conway for unpublished data, and Professor F. G. Smith, Dr A. N. Lasenby and others for many useful discussions.

Received 30 March; accepted 12 June 1982.

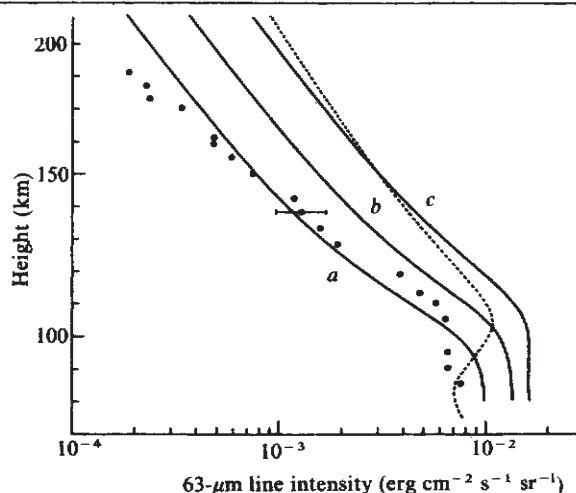
1. Raine, D. J. *Rep. Prog. Phys.* **44**, 1151-1195 (1981).
2. Ostvath, I. & Schücking, E. *Nature* **193**, 1168-1169 (1962).
3. Tabara, H. & Inoue, M. *Astr. Astrophys. Suppl.* **39**, 379-393 (1980).
4. Simard-Normadin, M., Kronberg, P. P. & Button, S. *Astrophys. J. Suppl.* **45**, 97-112 (1981).
5. Haves, P. *Mon. Not. R. astr. Soc.* **173**, 553-568 (1975).
6. Laing, R. A. *Mon. Not. R. astr. Soc.* **195**, 261-324 (1981).
7. Ekers, R. D. *Aust. J. Phys. Astrophys. Suppl.* **6**, 3-87 (1969).
8. Conway, R. G., Burn, B. J. & Vallée, J. P. *Astr. Astrophys. Suppl.* **27**, 155-170 (1977).
9. Simard-Normadin, M. & Kronberg, P. P. *Astrophys. J.* **242**, 79-94 (1980).
10. Högbom, J. A. *Astr. Astrophys. Suppl.* **36**, 173-192 (1979).
11. Laing, R. A. *Mon. Not. R. astr. Soc.* **193**, 439-449 (1980).
12. Vallée, J. P. *Nature* **254**, 23-26 (1975).
13. Partridge, R. B. *Phys. Scr.* **21**, 624-629 (1980).
14. Collins, C. B. & Hawking, S. W. *Mon. Not. R. astr. Soc.* **162**, 307-320 (1973).
15. Smoot, G. F. *Phys. Scr.* **21**, 619-623 (1980).
16. Gorenstein, M. V. & Smoot, G. F. *Astrophys. J.* **224**, 361-381 (1981).

## Thermospheric 63- $\mu\text{m}$ emission of atomic oxygen in local thermodynamic equilibrium

N. Iwagami & T. Ogawa

Geophysics Research Laboratory, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

The 63- $\mu\text{m}$  emission of atomic oxygen arising from the fine-structure transition ( $^3\text{P}_1$ - $^3\text{P}_2$ ) in the ground electronic state has been attributed to an important cooling mechanism of the thermosphere<sup>1</sup>; however, the zenith-emission intensities measured by Grossmann and Offermann<sup>2</sup> were much weaker than those expected from the theoretical calculation<sup>3</sup> based on local thermodynamic equilibrium (LTE). This discrepancy was interpreted by them to be due to a lower population temperature of  $\text{O}(^3\text{P}_j)$  ( $j=0, 1$  and  $2$ ) rather than LTE. If this is true, the 63- $\mu\text{m}$  emission is a much less important coolant of the thermosphere than has been believed. However, we have found that the formulation adopted by the previous calculation<sup>3</sup> is inadequate. Based on an improved formulation we have made a LTE model calculation for  $\text{O}(^3\text{P}_j)$  and found that calculated 63- $\mu\text{m}$  zenith intensities would be consistent with those measured.



**Fig. 1** Comparison of 63- $\mu\text{m}$  zenith intensities; our present calculation (solid curves *a*-*c*), the observation by Grossmann and Offermann<sup>2</sup> (●) and the previous calculation by Kockarts and Peetermans<sup>3</sup> (dotted curve).

In calculating the 63- $\mu\text{m}$  zenith intensity in the lower thermosphere below  $\sim 130$  km, the self-absorption effect should be considered, as the optical depth at the line centre exceeds unity at around 100 km. At the same time, the difference between emission and absorption line widths should be taken into account for the self-absorption process because of a non-isothermal thermosphere. The Doppler width of the 63- $\mu\text{m}$  line emitted in the upper thermosphere is considerably broader than that of the absorption line in the lower thermosphere. The formulation of ref. 3 considered the variation of the Doppler width with altitude (that is with temperature), but failed to treat adequately the difference between emission and absorption line profiles. In the present calculation the difference in the line profiles is formulated accurately, so that the wing portions of the emission line suffer much less self-absorption in the lower thermosphere than that evaluated from the formulation of ref. 3. The present formulation is similar to that of Strickland<sup>4</sup> for a UV dayglow, and is physically equivalent to that of Craig and Gille<sup>5</sup> for the 63- $\mu\text{m}$  emission.

The integrated line intensity in the zenith direction  $I(z)$  ( $\text{erg cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ ) can be calculated as

$$I(z) = \int_z^\infty \epsilon(z') T(z' \rightarrow z) dz' \quad (1)$$

where  $\epsilon$  ( $\text{erg cm}^{-3} \text{s}^{-1} \text{sr}^{-1}$ ) is the volume emissivity for the line, and  $T(z' \rightarrow z)$  is the transmission probability for the emission line which originates at  $z'$  and reaches  $z$ .  $T(z' \rightarrow z)$  can be expressed in terms of the normalized line profile function  $\phi_\nu$  ( $\text{Hz}^{-1}$ ) as

$$T(z' \rightarrow z) = \int_{\text{line}} \phi_\nu(z') \exp \left[ - \int_z^{z'} \kappa(z'') \phi_\nu(z'') dz'' \right] d\nu \quad (2)$$

where  $\kappa$  ( $\text{Hz cm}^{-1}$ ) is the absorption coefficient integrated over the line. The essential point of the present formulation is that the variation of  $\phi_\nu$  with altitude (that is with temperature) is exactly accounted for in evaluating equation (2).  $\epsilon$  and  $\kappa$  are obtained in the same way as in ref. 3; LTE is assumed for the  $\text{O}(^3\text{P}_j)$  distribution.

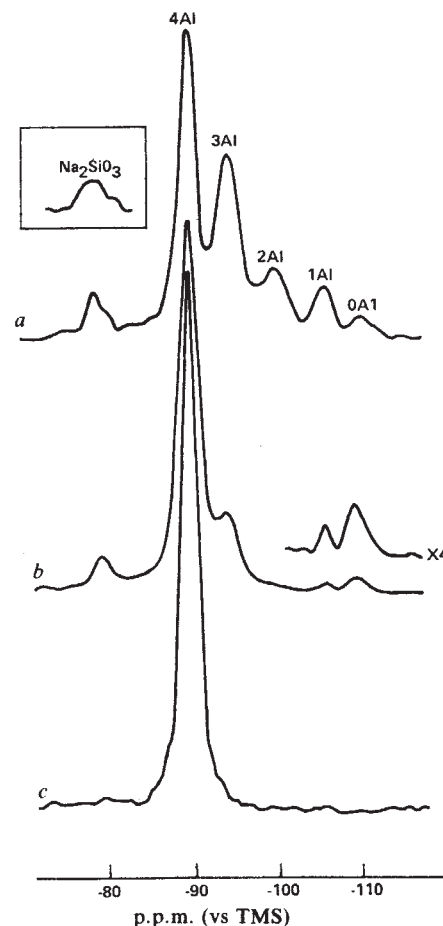
In Fig. 1 the present calculations using three atomic oxygen density profiles are compared with observation<sup>2</sup> and the previous calculation<sup>3</sup>. Curve *a* corresponds to the atomic oxygen densities which have a peak of  $2 \times 10^{11} \text{ cm}^{-3}$  at 96 km and decrease to  $7.5 \times 10^{10} \text{ cm}^{-3}$  at 110 km. The diffusive equilibrium for the atomic oxygen distribution is assumed in the altitudes above 110 km, where the temperature profile of CIRA 1972<sup>6</sup> is referred to. According to CIRA 1972, the thermopause temperature at the time of the observation is estimated to be 700 K from the 10.7-cm solar flux. The atomic oxygen densities

for curves *b* and *c* are larger by factors of 2 and 4, respectively, than those for curve *a*. The atomic oxygen densities for curve *c* are close to (only ~20% smaller than) those of CIRA 1972 at altitudes of 110 km and above. The atomic oxygen profile adopted in ref. 3 is close to that for curve *b* at around the peak altitudes, but is intermediate between those for curves *b* and *c* above 110 km, and is for a thermopause temperature of 750 K.

Curve *a* agrees well with measurements within experimental errors, although there remains a slight difference in intensity gradients. Note that the altitude profile of the observed intensity at around 100 km is better reproduced by curve *a* than by the previous calculation (dotted curve). Although the atomic oxygen densities for curve *a* are rather smaller than those estimated from CIRA 1972, the agreement between curve *a* and observation encourages us to adopt a LTE model. Unfortunately, we have a poor knowledge of the atomic oxygen profile for this particular observation made at night at high latitude. It is highly probable that the atomic oxygen densities at this observation would be considerably smaller than those of CIRA 1972. If we allow an uncertainty of a factor of 2 for the absolute calibration of the measurement, as inferred by the experimenters<sup>7</sup>, curve *b* might agree with the observation. Note that curve *b* was calculated for an atomic oxygen profile which is not far away from CIRA 1972. As a result, the LTE model for the  $O(^3P_J)$  levels does not seem to contradict the observed 63- $\mu\text{m}$  zenith intensity, but to reproduce it rather well.

Unless extremely small cross-sections are applicable to the collisional relaxations among the  $O(^3P_J)$  levels, LTE is a reasonable assumption for the  $O(^3P_J)$  levels as discussed previously<sup>8,9</sup>. Actually, one theory estimates the cross-sections for the mutual relaxation collisions to be between 0.01 and 1 of the cross-section for the gas kinetic collision<sup>10</sup>. Below 200 km the gas kinetic collision frequency is more than four orders of magnitude larger than the spontaneous transition probability of the  $^3P_1$  level. LTE distribution for the  $O(^3P_J)$  level is also inferred from the intensity ratio of the 1,302–1,305–1,306 Å triplet airglow<sup>11</sup>.

We conclude that there is no firm reason to favour a non-LTE for the population of the  $O(^3P_J)$  levels in the thermosphere.



**Fig. 1**  $^{29}\text{Si}$  NMR spectra at 39.5 MHz with magic angle spinning and proton dipolar decoupling for *a*, ZK4 with elemental composition  $\text{Si}/\text{Al} = 1.40$ , *b*, ZK4 with elemental composition  $\text{Si}/\text{Al} = 1.13$ , and *c*, zeolite A with  $\text{Si}/\text{Al} = 1.00$ . Assignment of NMR peaks to environments  $\text{Si}(n\text{Al})$ ,  $n = 4, 3, 2, 1, 0$  are indicated. Chemical shifts are referenced to tetramethylsilane (TMS). The  $^{29}\text{Si}$  NMR absorption for hydrated sodium metasilicate is shown in the inset.

To the extent that a given  $^{29}\text{Si}$  NMR absorption can be identified with a particular Si environment  $[\text{Si}(n\text{Al})]$ ,  $n = 0, 1, 2, 3, 4$ , the NMR data provide a direct measurement of the Al neighbour distribution. Previous  $^{29}\text{Si}$  NMR studies of zeolite A samples with  $\text{Si}/\text{Al}$  close to unity show single sharp absorptions with emical shifts in the range  $-88.3$  to  $-89.7$  p.p.m. versus tetramethylsilane<sup>1-4</sup>. A comparison of this chemical shift with the ranges observed for other aluminosilicates has led to the proposal that all silicon-oxygen tetrahedra in the zeolite A framework have three aluminium-oxygen tetrahedra and one silicon-oxygen tetrahedron as nearest neighbours, that is, that this single peak be assigned to the environment  $\text{Si}(3\text{Al})$  (refs 1-4). Such an arrangement necessarily involves Al-O-Al bonds, in violation of Loewenstein's rule<sup>9</sup>. Structural models for zeolite A have been proposed based on electron and neutron diffraction data<sup>2,3</sup>, which are consistent with both the observed 24 Å unit cell and this interpretation of the NMR data. However, there remains major disagreement in the literature<sup>10,11</sup>. Smith and Pluth have argued in favour of a Loewenstein distribution based on their X-ray study of single crystals of dehydrated zeolite A in sodium and potassium forms<sup>10</sup>, and suggested that distortions due to exchangeable cations could result in space group violations.

The assignment of the single  $^{29}\text{Si}$  resonance in NaA to a particular environment  $\text{Si}(n\text{Al})$  is ambiguous. This is in contrast to faujasite where a range of compositions is available within the same crystal framework. Here the assignment is clear

## The characterization of Si-Al ordering in A-type zeolite (ZK4) by $^{29}\text{Si}$ NMR

Michael T. Melchior, David E. W. Vaughan,  
Richard H. Jarman & Allan J. Jacobson

Exxon Research and Engineering Company, PO Box 121, Linden,  
New Jersey 07036, USA

**The study of the  $^{29}\text{Si}$  NMR spectra of two A-type zeolites (ZK4) with  $\text{Si}/\text{Al}$  ratios of 1.18 and 1.40 shows that the Si, Al distribution obeys Loewenstein's rule. The NMR absorption observed at  $-89.1$  p.p.m. for both compositions is assigned to  $\text{Si}(4\text{Al})$ . This result suggests that the single peak observed for NaA ( $\text{Si}/\text{Al} = 1$ ) should also be assigned to  $\text{Si}(4\text{Al})$  and not  $\text{Si}(3\text{Al})$  as recently proposed<sup>1-4</sup>.**

In aluminosilicate structures where each silicon has four nearest-neighbour silicon or aluminium atoms, the  $^{29}\text{Si}$  chemical shifts fall into five ranges which can be correlated with the number of aluminium neighbours surrounding a given silicon<sup>5-7</sup>.



because at some compositions all five resonances  $\text{Si}(n\text{Al})$ ,  $n = 0 \dots 4$  are observed. Furthermore, the evolution of the spectra with composition confirms that all of the assigned  $^{29}\text{Si}$  NMR absorptions arise from the faujasite framework<sup>8,12,13</sup>. We have used this approach to assign the single  $^{29}\text{Si}$  NMR absorption observed for NaA by a study of the zeolite ZK4 which has an A-type framework but  $\text{Si}/\text{Al} > 1$  (ref. 14). Analysis of the results has shown that the Si, Al distribution in ZK4 follows Loewenstein's rule and suggests that, contrary to the previous work<sup>1-4</sup>, the single peak in NaA should be assigned to  $\text{Si}(4\text{Al})$ .

The zeolite ZK4 can be prepared using adaptations of the method described by Kerr<sup>14</sup>. Sample I was prepared by first dissolving 8.25 g of a desiccant grade silica gel (Grace 923) in 175 cm<sup>3</sup> 25% tetramethylammonium (TMA) hydroxide solution by mixing at 25 °C for 6 days, then adding a solution of 10.75 g sodium aluminate (Baker) dissolved in 125 g H<sub>2</sub>O. After thorough homogenization the sample was aged for 20 h at 100 °C in a Teflon jar. Sample II was prepared by reacting a similar composition [8.0(TMA)<sub>2</sub>O: 1.45 Na<sub>2</sub>O: Al<sub>2</sub>O<sub>3</sub>: 5.33 SiO<sub>2</sub>: 320 H<sub>2</sub>O] made using a silica sol (Ludox HS40) at 100 °C for 24 h with rotation mixing. A reference sample of zeolite A (sample III) was made by reacting a sodium silicate composition of 2 Na<sub>2</sub>O: Al<sub>2</sub>O<sub>3</sub>: 2.1 SiO<sub>2</sub>: 80 H<sub>2</sub>O for 2 h at 100 °C. Chemical analyses were made using a Jarrel Ash inductively coupled plasma/atomic emission spectrometer. X-ray analysis using a Siemens D500 showed highly crystalline single phase zeolite A and ZK4 and scanning electron microscopy showed all samples to be cubes in the 1–2 µm range. Elemental analyses were as follows. Sample I, 0.88 Na<sub>2</sub>O: 0.12(TMA)<sub>2</sub>O: 1.00 Al<sub>2</sub>O<sub>3</sub>: 2.26 SiO<sub>2</sub>; sample II, 0.86 Na<sub>2</sub>O: 0.14(TMA)<sub>2</sub>O: 1.00 Al<sub>2</sub>O<sub>3</sub>: 2.80 SiO<sub>2</sub>; sample III, 1.00 Na<sub>2</sub>O: 1.00 Al<sub>2</sub>O<sub>3</sub>: 1.99 SiO<sub>2</sub>.

The  $^{29}\text{Si}$  NMR data were obtained at 39.5 MHz on a JEOL FX200WB spectrometer using the combined techniques of magic-angle sample spinning and proton dipolar decoupling. The NMR spectra are shown in Fig. 1. The spectrum obtained for NaA (sample III) is shown in Fig. 1c and consists of a single sharp absorption at -89.1 p.p.m., deshielded from tetramethylsilane (TMS). Figure 1a shows the  $^{29}\text{Si}$  NMR spectrum for sample II, ZK4 with elemental composition  $\text{Si}/\text{Al} = 1.40$ . This spectrum shows six distinct absorption peaks. One of these, at -77.8 p.p.m., is assigned to a polysilicate impurity, probably hydrated sodium metasilicate  $\text{Na}_2\text{SiO}_3$ , the  $^{29}\text{Si}$  NMR signal from which is shown in the inset. This impurity represents <3% of the total sample and is not detected in the X-ray diffraction measurements. The other five  $^{29}\text{Si}$  NMR absorptions at -89.1, -93.9, -99.5, -106.1, -110.7 p.p.m. are well within the range reported for four coordinate silicon<sup>2</sup> and are assigned to the five silicon environments  $\text{Si}(4\text{Al})$ ,  $\text{Si}(3\text{Al})$ ,  $\text{Si}(2\text{Al})$ ,  $\text{Si}(1\text{Al})$ ,  $\text{Si}(0\text{Al})$ , respectively, as indicated in Fig. 1. With the exception of the peak assigned to  $\text{Si}(4\text{Al})$ , the observed chemical shifts fall within the appropriate ranges for  $\text{Si}(n\text{Al})$  with  $n = 3, 2, 1, 0$  given in, for example, Klinowski *et al.*<sup>7</sup>.

To support this assignment we now show that it is consistent with the elemental composition and Loewenstein's rule. We have shown<sup>8</sup> that in zeolites which obey Loewenstein's rule the framework composition can be obtained from the first moment of the aluminium neighbour distribution, as reflected in the  $^{29}\text{Si}$  NMR spectrum. From the relative populations 0.317, 0.375, 0.168, 0.103 and 0.037 for  $\text{Si}(4\text{Al})$ , ...,  $\text{Si}(0\text{Al})$  we obtain the framework composition  $\text{Si}/\text{Al} = 1.41$ , in excellent agreement with the elemental analysis result. However, when the peak at -89.2 is assigned to  $\text{Si}(3\text{Al})$  and '3:1 ordering'<sup>3</sup> is assumed, the composition derived from the NMR data,  $\text{Si}/\text{Al} = 1.8$ , clearly does not agree with the elemental analysis, and the peak at -110.7 p.p.m. is not accounted for.

Additional corroboration of our assignment comes from the  $^{29}\text{Si}$  NMR spectrum of sample I, ZK4 with  $\text{Si}/\text{Al} = 1.13$ , shown in Fig. 1b. This spectrum shows major peaks at -89.1 and -93.9 assigned to  $\text{Si}(4\text{Al})$  and  $\text{Si}(3\text{Al})$  and weak absorptions at -99.5, -106.1 and -110.5 assigned to  $\text{Si}(2\text{Al})$ ,  $\text{Si}(1\text{Al})$  and  $\text{Si}(0\text{Al})$ . The relative intensities observed for  $\text{Si}(n\text{Al})$ ,  $n = 4$ ,

3, 2, 1, 0 are 0.642:0.246:0.036:0.026:0.051, from which we calculate a framework composition  $\text{Si}/\text{Al} = 1.18$ , again in agreement with elemental analysis.

Thus the  $^{29}\text{Si}$  NMR spectra for NaA and the two ZK4 samples shown in Fig. 1 exhibit a consistent and systematic variation constrained by Loewenstein's rule similar to that observed for NaX zeolites provided it is assumed that the single  $^{29}\text{Si}$  NMR absorption in NaA represents a  $\text{Si}(4\text{Al})$  environment. We will discuss the details of the  $^{29}\text{Si}$  NMR spectra of ZK4 over the range  $1 < \text{Si}/\text{Al} < 1.4$  elsewhere.

Received 10 March; accepted 12 May 1982.

- Engelhardt, G., Zeigan, D., Lippmaa, E. & Magi, M. *Z. Anorg. Allg. Chem.* **468**, 35–38 (1980).
- Thomas, J. M., Bursill, L. A., Lodge, E. A., Cheetham, A. K. & Fyfe, C. A. *JCS chem. Commun.* 276–277 (1981).
- Bursill, L. A., Lodge, E. A., Thomas, J. M. & Cheetham, A. K. *J. phys. Chem.* **85**, 2409–2421 (1981).
- Lippmaa, E., Magi, M., Samoson, A., Tarmak, M. & Engelhardt, G. *J. Am. chem. Soc.* **103**, 4992–4996 (1981).
- Lippmaa, E. T., Alla, M. A., Pehk, T. J. & Engelhardt, G. *J. Am. chem. Soc.* **100**, 1929–1930 (1978).
- Lippmaa, E., Magi, M., Samoson, A., Engelhardt, G. & Grimmer, A. R. *J. Am. chem. Soc.* **102**, 4889–4893 (1980).
- Klinowski, J., Thomas, J. M., Fyfe, C. A. & Hartman, J. S. *J. phys. Chem.* **85**, 2409–2421 (1981).
- Melchior, M. T., Vaughan, D. E. W. & Jacobson, A. J. *J. Am. chem. Soc.* (in the press).
- Loewenstein, W. *Am. Miner.* **39**, 92–96 (1954).
- Smith, J. V. & Pluth, J. J. *Nature* **291**, 265 (1981).
- Bursill, L. A., Lodge, E. A. & Thomas, J. M. *Nature* **291**, 265 (1981).
- Randas, S., Thomas, J. M., Klinowski, J., Fyfe, C. A. & Hartman, J. S. *Nature* **292**, 228 (1981).
- Engelhardt, E. *et al. Zeolites* **2**, 59 (1982).
- Kerr, G. T. *Inorg. Chem.* **9**, 1537–1539 (1966).

## A submarine hydrothermal manganese deposit from the south-west Pacific island arc

D. S. Cronan\*, G. P. Glasby†, S. A. Moorby\*, J. Thomson‡, K. E. Knedler§ & J. C. McDougall†

\* Department of Geology, Imperial College of Science and Technology, Prince Consort Road, London SW7 2BP, UK

† New Zealand Oceanographic Institute, PO Box 12-346, Wellington, New Zealand

‡ Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

§ Department of Geology, Victoria University of Wellington, Wellington, New Zealand

Hydrothermal deposits of a wide variety of types are being found with increasing frequency on or near actively spreading mid-ocean ridges<sup>1-4</sup>. However, they also have a potential to occur in other submarine volcanic settings, including island arcs<sup>5</sup>. To follow up indications of mineralization associated with submarine hydrothermal activity in the south-west Pacific island arc<sup>6</sup>, a joint New Zealand Oceanographic Institute/Imperial College research cruise was mounted in May 1981 aboard the RV *Tangaroa*. During this cruise, over 130 sampling stations were occupied, at one of which were dredged manganese deposits with strong hydrothermal affinities. This is the first report of such deposits from an island arc setting.

The manganese deposits occur in a small area on the flanks of the Tonga–Kermadec Ridge at 26°56'S, 177°37'W, in a water depth of ~1,700 m. The Tonga–Kermadec Ridge is a volcanically active feature stretching from Tonga to New Zealand, and constitutes the overriding edge of the Austral–Indian Plate. In appearance, the deposits consist of pieces of layered and laminated manganese oxides with a sub-metallic lustre similar to the hydrothermal manganese deposits found in the TAG hydrothermal field on the Mid-Atlantic Ridge, and more recently in the Galapagos Rift<sup>7</sup>. Of the numerous pieces of Mn-oxide crust recovered, the two samples described here are very different from each other in appearance. Sample U23/1 (Fig. 1a) shows features typical of many of the recovered fragments. It consists

**Table 1** Bulk chemical composition of Mn-oxide crusts from the south-west Pacific island arc

| Sample                                                              | Mn (%) | Fe (%) | Ca (%) | Al (%) | Co (p.p.m.) | Ni (p.p.m.) | Cu (p.p.m.) | Zn (p.p.m.) |
|---------------------------------------------------------------------|--------|--------|--------|--------|-------------|-------------|-------------|-------------|
| U23/1                                                               |        |        |        |        |             |             |             |             |
| Surface granular layer                                              | 52     | 0.24   | 2.1    | 0.25   | <3          | 320         | 100         | 490         |
| Surface-4 mm                                                        | 49     | 0.05   | 2.2    | 0.03   | 8           | 93          | 24          | 110         |
| 4-10 mm                                                             | 50     | 0.06   | 1.6    | 0.07   | <3          | 150         | 26          | 170         |
| 10-15 mm                                                            | 49     | 0.01   | 1.9    | 0.02   | <3          | 15          | 34          | 29          |
| 15-20 mm                                                            | 49     | 0.02   | 1.8    | 0.03   | <3          | 40          | 34          | 50          |
| 20-27 mm                                                            | 52     | 0.08   | 1.8    | 0.08   | <3          | 105         | 29          | 150         |
| Basal granular material                                             | 46     | 0.32   | 1.9    | 0.28   | 12          | 480         | 170         | 710         |
| U23/2                                                               |        |        |        |        |             |             |             |             |
| Granular crust                                                      | 38     | 4.5    | 2.1    | 1.1    | 55          | 730         | 200         | 760         |
| Sediment substrate                                                  | 0.25   | 3.7    | 2.6    | 15.6   | —           | 20          | 28          | 90          |
| TAG area <sup>17</sup>                                              | 39     | 0.06   | —      | —      | 19          | 350         | 43          | —           |
| Famous area <sup>1</sup>                                            | 28     | 10.5   | 2.9    | 0.6    | 82          | 370         | 206         | 83          |
| Galapagos Mounds <sup>2</sup>                                       | 50     | 0.26   | 1.5    | 0.19   | 5           | 470         | 100         | 380         |
| Hydrogenous Mn-oxide deposits (Pacific Ocean average) <sup>16</sup> | 20     | 12     | 2.0    | 3.1    | 3400        | 6300        | 3900        | 680         |
| Diagenetic Mn-oxide deposits <sup>16</sup>                          | 34     | 1.6    | —      | —      | 75          | 970         | 650         | —           |

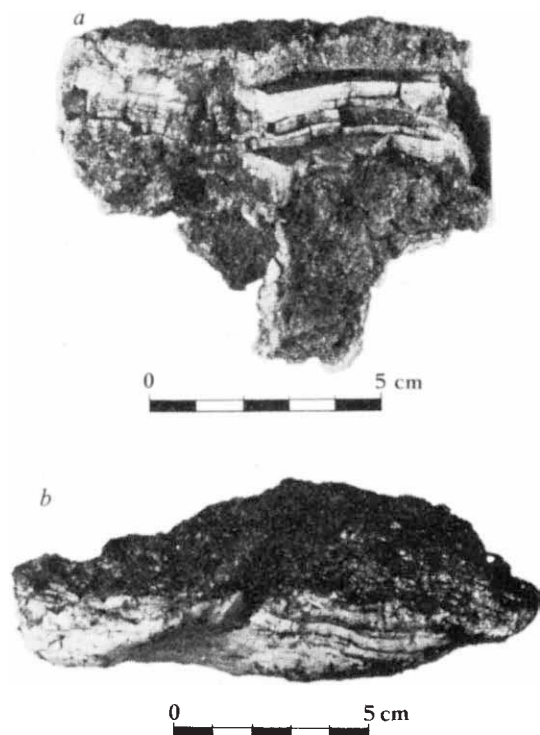
of a surface granular layer 0–2 mm thick beneath which is 20–25 mm of very hard, purplish-black crust showing many internal laminations 1–5 mm thick. These very hard layers are quite different from the hydrogenous ferromanganese oxide crusts commonly encountered on exposed rock surfaces on the deep ocean floor<sup>8</sup>. Beneath them is a basal layer 20–40 mm thick comprising massive black crust showing no internal features but having a microgranular surface texture and very irregular shape. The second sample, U23/2 (Fig. 1b) is unique in having an attached substrate. The crust itself is up to 25 mm thick, has a very irregular, granular surface and no visible internal features. The substrate is a minimum of 20 mm thick and consists of indurated clay-rich sediment which appears to be penetrated in places by the overlying crust.

Sub-samples of each of the two crusts were analysed by atomic absorption spectrophotometry and X-ray diffrac-

tometry. Mineralogically, the crusts consist mainly of well-crystalline birnessite. However, sample U23/2 and the surface layer of U23/1 contain traces of todorokite. Hydrothermal manganese deposits from on and near mid-ocean ridges have been reported as containing birnessite, as well as todorokite<sup>1,2,4</sup>. In this respect the deposits are, in large part, quite different from normal hydrogenous ferromanganese oxide encrustations which typically consist mainly of  $\delta$ -MnO<sub>2</sub> or amorphous ferromanganese oxides, sometimes with subordinate todorokite<sup>8</sup>. Thus mineralogical evidence is in accord with a non-hydrogenous origin for the crusts.

The results of the chemical analyses of the samples are given in Table 1, together with data on other marine ferromanganese oxide deposits for comparison. The crusts are very rich in Mn, up to a maximum of ~50%, and are very low in Fe. In this respect they are similar to hydrothermal manganese crusts from the TAG hydrothermal field and certain other submarine hydrothermal localities (Table 1). Calcium is low in all the samples, and Al lower. The minor elements Cu, Ni, Co and Zn are generally lower than their concentrations in hydrogenous ferromanganese oxides (Table 1), providing further evidence for the non-hydrogenous origin of the deposits. Only Zn approaches the value normally found in hydrogenous manganese crusts.

Whereas chemically, the deposits described here are quite different from hydrogenous ferromanganese oxide deposits, being characterized by extremely high Mn/Fe ratios and very low trace metal contents, the granular type material does have a higher trace metal content and shows similarities with diagenetically formed ferromanganese oxide deposits formed by the upward remobilization of manganese through buried sediments (Table 1). However, sample U23/2 has an indurated sediment substrate through which migration of metals is unlikely to have occurred. Furthermore, the environmental setting of the deposits is quite different from that of diagenetic ferromanganese oxide deposits. They occur not in a continental margin situation underlain by rapidly deposited sediments which are reducing at shallow depth, as are many diagenetically formed ferromanganese oxide deposits, but on an elevated volcanic ridge far from a continental source of sediment. Thus, in spite of some similarities with diagenetically formed deposits, we conclude that the crusts described here are not diagenetic in origin but that their overall composition reflects a hydrothermal origin. The slightly elevated concentrations of Fe, Al and certain minor elements in the granular material may reflect an increased hydrogenous influence on those parts of the crust. They may be transitional between pure hydrothermal and pure hydrogenous ferromanganese oxides and represent deposition during a period of reduced hydrothermal activity.



**Fig. 1** a, Side view of sample U23/1. Surface of crust is at top.  
b, Sample U23/2.



**Table 2** Isotopic data on manganese oxide crust U23/1 from the south-west Pacific island arc

| Depth<br>(mm) | U<br>(p.p.m.) | Th<br>(p.p.m.) | U/Th    | $^{234}\text{U}/^{238}\text{U}$<br>activity<br>ratio | $^{230}\text{Th}/^{232}\text{U}$<br>activity<br>ratio | $^{234}\text{U}$<br>(d.p.m. g <sup>-1</sup> ) | $^{230}\text{Th}$<br>(d.p.m. g <sup>-1</sup> ) | $^{230}\text{Th}/^{234}\text{U}$<br>activity<br>ratio | Age*<br>(kyr) |
|---------------|---------------|----------------|---------|------------------------------------------------------|-------------------------------------------------------|-----------------------------------------------|------------------------------------------------|-------------------------------------------------------|---------------|
| Surface       | 6.88±0.28     | 0.25±0.10      | 28±11   | 1.09±0.06                                            | 94±36                                                 | 5.59±0.22                                     | 5.79±0.24                                      | 1.04±0.06                                             | —             |
| 0-4           | 8.97±0.23     | 0.71±0.13      | 13±2    | 1.17±0.03                                            | 16±3                                                  | 7.86±0.19                                     | 2.77±0.14                                      | 0.35±0.02                                             | 46.2±3.4      |
| 4-10          | 4.83±0.15     | 0.31±0.06      | 16±3    | 1.24±0.05                                            | 11±2                                                  | 4.47±0.13                                     | 0.83±0.05                                      | 0.19±0.01                                             | 22.7±1.4      |
| 10-15         | 13.9±0.4      | 0.96±0.12      | 14±2    | 1.09±0.03                                            | 8.5±1.1                                               | 11.33±0.31                                    | 1.98±0.09                                      | 0.17±0.01                                             | 20.2±1.3      |
| 15-20         | 4.89±0.20     | 0.75±0.10      | 6.5±0.9 | 1.06±0.05                                            | 6.2±0.9                                               | 3.87±0.16                                     | 1.14±0.06                                      | 0.29±0.02                                             | 37.1±3.2      |
| 20-27         | 3.69±0.11     | 0.38±0.06      | 9.7±1.6 | 1.10±0.04                                            | 13±2                                                  | 3.03±0.09                                     | 1.15±0.05                                      | 0.38±0.02                                             | 51.5±3.6      |
| 27-           | 3.77±0.13     | 1.6±0.2        | 2.4±0.3 | 1.20±0.05                                            | 14±2                                                  | 3.37±0.11                                     | 5.43±0.20                                      | 1.61±0.07                                             | —             |

\* Ages calculated from ingrowth of  $^{230}\text{Th}$  towards radioactive equilibrium with  $^{234}\text{U}$ , where initial  $^{230}\text{Th}$  is assumed to be zero.

To verify the hydrothermal nature of the deposits, radiochemical analysis for uranium series isotopes was undertaken. Data for a profile of samples from crust U23/1, obtained by  $\alpha$  spectrometry with isotopic tracers, are presented in Table 2. The relatively high uranium and low thorium values found are characteristic of the values found by Bonatti and co-workers<sup>9,10</sup> for hydrothermal metalliferous deposits, as are the  $^{234}\text{U}/^{238}\text{U}$  activity ratios.

Derivation of accumulation rates from  $^{230}\text{Th}/^{234}\text{U}$  data for the type of material being described here is not unequivocal, due to uncertainties on the precise initial conditions<sup>11-13</sup>. In Table 2, ages are listed for the layered section of crust U23/1, calculated as a function of time from the experimental  $^{234}\text{U}/^{238}\text{U}$  and  $^{230}\text{Th}/^{234}\text{U}$  activity ratios<sup>14</sup>, on the assumption that the crust is a chemically closed system and that  $^{230}\text{Th}/^{234}\text{U}$  is zero at the time of formation. This latter assumption is probably incorrect as it is certain that the small content of  $^{232}\text{Th}$  was accompanied by some  $^{230}\text{Th}$ : the consequence is that the calculated ages in Table 2 are overestimated and accumulation rates derived from them underestimated. It is evident that the calculated ages cannot be entirely accurate from the fact that the data in Table 2 indicate an older age for the near-surface layer of U23/1 than for immediately underlying layers. This may be due to contamination from the surface sample, which has a high  $^{230}\text{Th}/^{234}\text{U}$  ratio. Nevertheless, the data are sufficient for us to conclude that the layered section of crust U23/1 accumulated at a rate of at least 500 mm Myr<sup>-1</sup>, far higher than the accumulation rates of most hydrogenous manganese nodules studied to date<sup>15</sup>, and consistent with it being hydrothermal in origin.

These data represent the first report of a submarine hydrothermal manganese deposit in the south-west Pacific island arc. However, it is not from a back-arc spreading centre as might be expected in view of the common occurrence of hydrothermal manganese deposits at mid-ocean ridge spreading centres, but from the volcanic line itself. It is clear from better investigated submarine hydrothermal systems such as those on oceanic hydrothermal centres in the Red Sea and Galapagos Rift that manganese deposits of the type described here from the south-west Pacific are the last phases to precipitate in a hydrothermal fractionation sequence, earlier members of which include iron silicates and polymetallic sulphides of possible economic value<sup>8</sup>. It is possible therefore that polymetallic sulphide deposits may occur on the sea floor in the vicinity of the manganese deposit that we have found.

It is evident from the discovery of a submarine hydrothermal manganese deposit in an island arc setting that the area of sea floor on which submarine hydrothermal deposits can now be expected must be expanded to include island arcs as well as mid-ocean ridge spreading centres. It is clear therefore that in addition to mid-ocean ridges, modern island arcs must also be targets in future exploration for potentially exploitable submarine massive sulphide deposits.

We thank the master and crew of the RV *Tangaroa* for assistance in the collection of the samples. Financial support was provided by the Natural Environment Research Council (UK) and the Department of Scientific and Industrial Research (New Zealand).

Received 24 March; accepted 11 June 1982.

- Hoffert, M. *et al. Oceanol. Acta* **1**, 73 (1978).
- Corliss, J. B., Lyle, M., Dymond, J. & Crane, K. *Earth planet. Sci. Lett.* **40**, 12 (1978).
- Francheteau, J. *et al. Nature* **227**, 523 (1979).
- Lonsdale, P., Burns, V. M. & Fisk, M. J. *Geol.* **88**, 611 (1980).
- Cronan, D. S. *Nature* **262**, 567 (1976).
- Cronan, D. S. & Thompson, B. *Trans. Instn Min. Metal.* **B87**, 87 (1978).
- Malahoff, A., Cronan, D. S., Skirrow, R. & Emblin, R. *Mar. Min.* (in the press).
- Cronan, D. S. *Underwater Minerals* (Academic, London, 1980).
- Bonatti, E., Kraemer, T. & Rydell, H. in *Ferromanganese Deposits on the Ocean Floor* (ed. Horn, D.) 149 (NSF, Washington DC, 1972).
- Bonatti, E. *Ann. Rev. Earth planet. Sci.* **3**, 401 (1975).
- Moore, W. S. & Vogt, P. R. *Earth planet. Sci. Lett.* **29**, 349 (1976).
- Lalou, C., Bricchet, E., Ku, T. L. & Jehanno, C. *Mar. Geol.* **24**, 245 (1977).
- Burnett, W. C. & Piper, D. Z. *Nature* **265**, 569 (1977).
- Kaufmann, A. & Broecker, W. S. *J. geophys. Res.* **70**, 4039 (1965).
- Ku, T. L. in *Marine Manganese Deposits* (ed. Glasby, G.) 249 (Elsevier, Amsterdam, 1977).
- Cronan, D. S. in *Chemical Oceanography* Vol. 5 (eds Riley, J. & Chester, R.) 217 (Wiley, New York, 1976).
- Scott, M. R., Scott, R. B., Rona, P. A., Butler, L. W. & Nalwalk, A. J. *Geophys. Res. Lett.* **1**, 355 (1974).

## Linkage of North Atlantic and Southern Ocean deep-water circulation during glacial intervals

Bruce H. Corliss

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

**The Antarctic Circumpolar Current (ACC) is one of the major current systems in the world ocean, affecting circulation in all of the major ocean basins. The ACC flows eastwards around Antarctica from the surface to ~4,000 m depth with a transport on the order of 125 Sverdrups<sup>1</sup>. Deep-sea benthic foraminifera from one piston core (E49-18) from the south-east Indian Ocean sector of the Southern Ocean were analysed to determine deep-water-mass-circulation conditions within the ACC during the late Quaternary. I report here that benthic foraminiferal faunal patterns closely match Northern Hemisphere ice-volume changes shown by planktonic foraminiferal oxygen-isotopic stratigraphy. The glacial-interglacial faunal oscillations are interpreted to reflect the presence of a deep-water mass in the ACC during glacial times which differs from modern Circumpolar Deep Water. I suggest that deep-water circulation changes within the ACC during the interval 440,000 to the present are directly linked to changes in North Atlantic Deep Water (NADW) circulation.**

Recent studies of the distribution of deep-sea benthic foraminifera from trigger core-tops have demonstrated a correlation between benthic foraminiferal faunal assemblages and deep-water and abyssal water masses throughout the oceans<sup>2-9</sup>. The correlation between faunal assemblages and water masses provides a means by which Quaternary bottom-water circulation conditions can be determined from the analysis of piston core samples. The presence of a particular faunal assemblage in a piston core sample can be used to infer the presence of the associated water mass at the time the sediment was deposited.

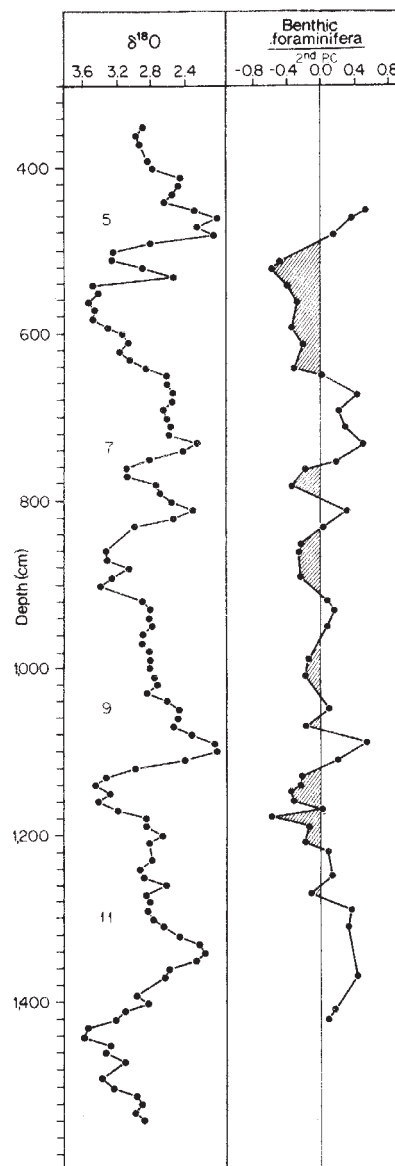


Core E49-18 was taken at 46°03.0' S, 90°09.3' E, 3,253 m water depth on the southern flank of the South-east Indian Ridge. This core contains foraminiferal and diatom oozes with radiolarians present, and has calcium carbonate percentages ranging from 15 to 93% (ref. 10). A mixture of Circumpolar Deep Water and Antarctic Bottom Water overlies this core, flowing eastward within the ACC with bottom-water potential temperatures of 0.7°C, salinity of 34.72‰, and dissolved oxygen contents of 4.9 m<sup>-1</sup> (ref. 11). Oxygen isotope stratigraphy presented previously<sup>10</sup> (Fig. 1) shows the presence of isotopic stages 12-5 (~440,000-128,000 yr BP). A recognizable isotopic stratigraphy was not found in the upper 4 m of the core<sup>10</sup> and this interval is not considered here. Two biostratigraphical datum levels, the extinctions of the coccolith *Pseudoemiliania lacunosa* at 458,000 yr BP (ref. 12) and the radiolarian *Stylatractus universus* at 425,000 yr BP (ref. 13) were used to identify isotopic stage 12 and confirm the placement of the isotopic stages<sup>10</sup>. Approximately 100 benthic foraminifera were picked from the >150-µm fraction, and identified and counted following the taxonomy outlined in ref. 14. Principal component analysis<sup>15</sup> of the faunal data is used to identify faunal assemblages. The first principal component represents an average of the faunal data, and the second principal component is used to identify the faunal assemblages<sup>4</sup>. A more complete discussion of the E49-18 data and data from three additional piston cores beneath the ACC are presented elsewhere<sup>16</sup>.

The second principal component identifies two faunal assemblages, with the benthic foraminiferal faunal oscillations following the oxygen isotope stratigraphy quite closely. The first assemblage (positive loadings) is dominated by *Globocassidulina subglobosa* (Brady) and a mixture of species including *Ehrenbergina trigona* Göes, *Planulina wuellerstorfi* (Schwager), and *Gyroidinoides soldanii* (d'Orbigny), and is found generally during the warm interglacial isotopic stages 11, 9, 7, and 5. The second assemblage (negative loadings) reflects the dominance of *Melonis barleeanum* (Williamson), *Melonis pompilioides* (Fichtel and Moll) and *Uvigerina peregrina* Cushman. This assemblage is found generally during glacial isotopic stages 10, 8, and 6, and intermediate cool intervals including the latter part of stage 11 and the middle of stages 9 and 7.

The *G. subglobosa* assemblage (positive loadings) is similar to a faunal assemblage found in the surface sediments of the south-east Indian Ocean associated with a mixture of Antarctic Bottom Water and Circumpolar Deep Water with potential temperatures of 0.6-0.8°C. The presence of the *G. subglobosa* assemblage suggests that this type of water was present at this location during interglacial times. Neither the *Melonis-Uvigerina* assemblage nor the high abundances of *M. barleeanum* (up to 28%) has been previously documented in surface sediments beneath the ACC<sup>5,14</sup>. Hence, the presence of a *Melonis-Uvigerina* assemblage indicates that a water mass was present within the ACC during glacial times within the 440,000-128,000 yr BP interval which was distinctly different from the present-day Circumpolar Deep Water. Faunal data<sup>14</sup> from E49-23 taken ~400 km near E49-18 show that these glacial-interglacial faunal patterns existed through isotopic stages 5-1 (128,000 yr BP to the present) in this area.

The composition of the faunas and timing of the faunal changes in E49-18 from the Southern Ocean are similar to those found by Streeter and Shackleton<sup>17</sup> in V29-179 (44°00.6' N, 24°32.4' W) from the north-east Atlantic. During glacial times within the past 125,000 yr, a *Uvigerina-M. barleeanum*-*Astrononion* assemblage was present. *Uvigerina* dominated the cool intervals of stage 5 and in stages 4 and 3, with *M. barleeanum* and *Astrononion* increasing in abundance in the latter part of stage 3 and in stage 2. The dominance of *Uvigerina* during glacial times, which was noted throughout the North Atlantic<sup>17-19</sup>, was suggested to reflect the reduction or elimination of the young, well-oxygenated North Atlantic Deep Water, resulting from the cooling and stratification of the Norwegian Sea surface<sup>20</sup>. Independent geochemical evidence<sup>21,22</sup> supports the hypothesis of elimination or reduc-



**Fig. 1** Benthic foraminiferal data and planktonic foraminiferal oxygen isotope stratigraphy<sup>10</sup> during the interval ~440,000-128,000 yr BP for E49-18, which is located on the southern flank of the south-east Indian Ridge in the Indian Ocean sector of the Southern Ocean. The second principal component (2nd PC) identifies two benthic foraminiferal faunal assemblages. One assemblage (negative loadings) indicates the dominance of *Melonis barleeanum*, *Melonis pompilioides*, and *Uvigerina peregrina* and is found during glacial intervals and intermediate cool intervals. A second assemblage (positive loadings) indicates the dominance of *Globocassidulina subglobosa*, and a mixture of species including *Ehrenbergina trigona*, *Planulina wuellerstorfi*, and *Gyroidinoides soldanii*, and is found generally during warm interglacial intervals.

tion of NADW during glacial intervals. The similarity of the faunal record from the north-east Atlantic and the Southern Ocean suggest that the circulation patterns are related, with the deep-water circulation changes in the Southern Ocean and the North Atlantic directly linked.

Circulation in the deep North Atlantic is dominated by NADW which forms as a result of dense Norwegian Sea water flowing over the sills of the Denmark Straits and the Iceland-Faeroe Ridge. This water mixes with Labrador Sea Water and Mediterranean Water<sup>23</sup> as it flows south in the Atlantic to the Southern Ocean where it flows into and becomes an important component of Circumpolar Deep Water<sup>24</sup>. The NADW has warm temperatures, high salinities, and high dissolved oxygen

contents relative to the Circumpolar Deep Water<sup>24</sup>. The reduction or cessation of NADW during glacial times would likely change the hydrographic characteristics of the deep water within the ACC, creating a Glacial Circumpolar Deep Water reflected by the *Melonis-Uvigerina* faunal assemblage. It seems unlikely that the Glacial Circumpolar Deep Water was a young, well-oxygenated water mass formed in the Southern Ocean, since the CaCO<sub>3</sub> record from E49-18 (ref. 10) shows glacial intervals to have low amounts of CaCO<sub>3</sub> relative to the interglacial intervals. If deep-water formation was taking place, a young, well-oxygenated water mass would enhance CaCO<sub>3</sub> preservation. Thus, the Glacial Circumpolar Deep Water was probably a modification of the present-day Circumpolar Deep Water.

A change in deep-water characteristics within the ACC probably affected the Southern Ocean heat budget. That a change in the Southern Ocean heat budget did occur is shown by the migration of the Polar Front<sup>25-27</sup> and by the change in the extent of summer sea ice during glacial intervals. Summer sea ice during the last glacial interval is inferred to have extended to 55° S (ref. 28) in contrast to the present day when summer sea ice melts back to the Antarctic continent<sup>29</sup>. The lack of NADW in the ACC may have decreased deep-water temperatures, as the NADW is a source of warm water for the ACC in the present day. This cooling of deep water temperatures would have enhanced sea-ice formation and retarded summer sea-ice melting, because the warm Circumpolar Deep Water is considered important in providing heat for the melting of sea ice<sup>30</sup>.

The close correspondence between the bottom-water mass changes in the Southern Ocean and the Northern Hemisphere ice volume record suggests that bottom-water production in the Norwegian Sea and the amount of Northern Hemisphere ice are controlled by the same variables. For example, polar cooling would influence both the amount of Northern Hemisphere ice and the development of sea ice in the Norwegian Sea. An additional factor which may have influenced NADW production is Mediterranean Sea outflow water<sup>31</sup>, which flows westward from the Straits of Gibraltar<sup>23</sup>. Reid<sup>31</sup> has suggested that Mediterranean outflow water also flows northwards into the Norwegian Sea, providing high-salinity water for the formation of Norwegian Sea overflow water. It was speculated that the elimination of Mediterranean outflow water caused by tectonic events at the sill or a drop in sea level closing the Straits of Gibraltar would preclude the formation of NADW, or alter its composition<sup>31</sup>. Following this argument, I speculate that the amount of Mediterranean outflow water during the Quaternary may be directly controlled by sea-level changes associated with Northern Hemisphere ice volume changes. For example, a lowering of sea-level, caused by a build-up of Northern Hemisphere ice, would decrease but would not eliminate the amount of Mediterranean outflow water to flow into the Norwegian Sea, which in turn would decrease the amount of Norwegian Sea overflow water produced.

I conclude that the benthic foraminiferal data from E49-18 show distinct glacial-interglacial oscillations which are interpreted to reflect significant changes in the Circumpolar Deep Water during the late Quaternary. The benthic foraminiferal faunal patterns in the Southern Ocean are similar to those found in the North Atlantic, and I suggest that the water mass changes within the ACC are a direct result of the reduction or elimination of NADW during glacial intervals. NADW production may be influenced not only by polar cooling, but also by sea-level regulated production of Mediterranean outflow water.

I thank E. A. Boyle, M. L. Bremer, C. D. Hollister, S. Honjo, L. D. Keigwin Jr, M. S. McCartney, and J. J. Morley for discussions; G. P. Lohmann for assistance with the principal component analysis; B. U. Haq, D. A. Johnson and L. D. Keigwin Jr for reviewing the manuscript; Dennis Cassidy for assistance in obtaining *Eltanin* samples; Stefan E. Masse for drafting and Frank Mederios for photographing the figures; E. Evans for typing the manuscript and Marleen H. Jeglinski for technical assistance. This study was supported by NSF grant DPP-7821105 and DPP-8020479 (Division of Polar Programs).

This is Woods Hole Oceanographic Institution contribution no. 5118.

Received 15 March; accepted 28 May 1982.

1. Fandry, C. & Pillsbury, R. D. *J. phys. Oceanogr.* **9**, 449-455 (1979).
2. Streeter, S. S. *Quat. Res.* **3**, 131-141 (1973).
3. Schnitker, D. *Nature* **248**, 385-387 (1974).
4. Lohmann, G. P. *J. Foram. Res.* **8**, 6-34 (1978).
5. Corliss, B. H. *Mar. Geol.* **31**, 115-138 (1979).
6. Douglas, R. G. & Heitman, H. L. *Soc. Econ. Pal. Min., Spec. Publ.* **27**, 231-246 (1979).
7. Buzas, M. A. & Culver, S. J. *Science* **209**, 687-689 (1980).
8. Culver, S. J. & Buzas, M. A. *Nature* **290**, 328-329 (1981).
9. Bremer, M. L. & Lohmann, G. P. *Deep-Sea Res.* (in the press).
10. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
11. Rodman, M. thesis, Columbia Univ. (1977).
12. Thierstein, H. R., Geitzenauer, K. R., Molino, B. & Shackleton, N. J. *Geology* **5**, 400-404 (1979).
13. Morley, J. J. & Shackleton, N. J. *Geology* **6**, 309-311 (1978).
14. Corliss, B. H. *Micropaleontology* **25**, 1-19 (1979).
15. Klován, J. E. & Imbrie, J. *Int. Ass. math. Geol. J.* **3**, 61-78 (1971).
16. Corliss, B. H. *Deep-Sea Res.* (in the press).
17. Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168-171 (1979).
18. Streeter, S. S. *Geol. Soc. Am. Abstr.* **9**, 1192 (1977).
19. Schnitker, D. *Mar. Micropaleont.* **4**, 265-280 (1979).
20. Kellogg, T. B. *Boreas* **9**, 115-137 (1980).
21. Boyle, E. A. & Keigwin, L. D. Jr *EOS* **63**, 71-72 (1982).
22. Curry, W. B. & Lohmann, G. P. *EOS* **62**, 904 (1982).
23. Worthington, L. V. *On the North Atlantic Circulation* (The Johns Hopkins University Press, 1976).
24. Reid, J. L., Nowlin, W. D. Jr & Patzert, W. C. *J. phys. Oceanogr.* **7**, 62-91 (1977).
25. Climap Project Members *Science* **191**, 1131-1137 (1976).
26. Williams, D. F. *Mar. Micropaleont.* **1**, 363-375 (1976).
27. Morley, J. J. & Hays, J. D. *Quat. Res.* **12**, 396-408 (1979).
28. Hays, J. D., Lozano, J. A., Shackleton, N. J. & Irving, G. *Geol. Soc. Am. Mem.* **145**, 337-372 (1976).
29. Fletcher, J. O. in *National Science Foundation Memorandum* R17-3793 (Santa Monica, 1969).
30. Gordon, A. L. *J. geophys. Res.* **86**, 4193-4197 (1981).
31. Reid, J. G. *Deep-Sea Res.* **26A**, 1199-1223 (1979).

## Historical evidence for a dramatic increase in the nitrate component of acid rain

P. Brimblecombe

School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK

D. H. Stedman

Atmospheric and Ocean Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA

**The concentrations of nitrate and ammonium ion in rainfall have an important effect on precipitation pH. We have collected sets of rainfall analyses from non-urban sites in North America and western Europe, which start from last century, and we show here that there is a marked increase in the annual deposit of nitrate ion compared with relatively stable levels of ammonium ion deposit. The increase apparent in the data from the US parallels the increases in nitrogen oxide emissions from combustion processes and is large enough for the nitrate ion to contribute almost as much to the acidity of rainfall as the sulphate ion.**

The best source of early rainfall analyses proves to be the records of agricultural stations concerned with the importance of meteoric nitrogen to crop yields. In the UK analyses were made at Rothamsted as early as 1853, although NO<sub>3</sub> and NH<sub>3</sub> were not determined simultaneously until several years later. In North America the earliest analyses appear in the 1880s. It is important to acknowledge the difficulties in assembling long records from sources widely separated by time and geography. There is some internal evidence in the Rothamsted data to suggest that even early analyses can be quite reliable<sup>1</sup>, and as large changes are to be expected they should be evident even in quite noisy data. An increase in the concentration of the nitrate ion in rainfall was suspected as long ago as 1918, and

the principal source assumed to be combustion processes<sup>2</sup>. Ammonia, on the other hand, has long been considered to emanate largely from the soil<sup>3</sup> and today the industrial inputs are relatively small<sup>4</sup>. Hence the annual deposit of the ammonium ion should have remained reasonably constant over the past 100 yr. Miller<sup>5</sup> tabulated some of the earlier data, collected from sites all over the world. As can be seen in Fig. 1, the relative proportion of nitrogen as ammonia remains fairly constant, independent of locality and amount deposited. This pre-industrial value for the fraction of inorganic nitrogen present as the ammonium ion,  $\text{NH}_4^+ / (\text{NH}_4^+ + \text{NO}_3^-)$ , of  $\sim 0.7$  seems much reduced in present day North America and Europe, where values as low as half this are found. Such a change in the ratio would be expected if the nitrate ion concentration of precipitation were increasing and that of the ammonium ion remaining fairly constant.

Analytical records available for sites in eastern North America (including only one site outside the US, in Ottawa) and England and Benelux now cover a century, so annual deposits may be plotted as a function of time. Although analyses of rainfall composition have become more frequent recently, they are now often event or wet-only samples. Such results are

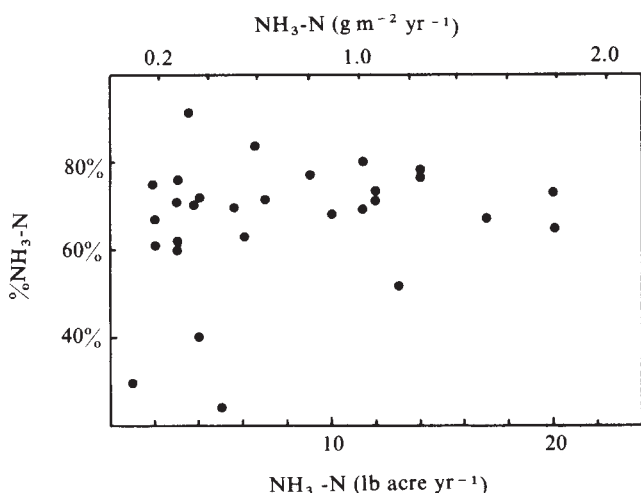


Fig. 1 The proportion of inorganic nitrogen deposited as ammonia as a function of ammonium ion deposit at sites in operation last century<sup>5</sup>.

difficult to compare with the older monthly samples which tend to show larger deposits (see ref. 6). Variation in the deposition of both nitrate and ammonium ion in the two regions are shown on semi-logarithmic axes in Figs 2 and 3. The annual values here are typically sums of 12-monthly measurements. Gaps in the data for the period 1920–50 show a decline of interest in rainfall chemistry among agricultural researchers. In both regions the deposit of nitrate ion increases throughout the duration of the record, while the amount of deposited ammonium ion is stable. The deposit of nitrate–nitrogen in eastern North America may be compared with the combustion source strength of the oxides of nitrogen over the same period. The early  $\text{NO}_x$  emissions for the US were determined from the amounts of energy obtained from coal, wood, gas and petroleum combustion<sup>7</sup>, weighted for yields of  $\text{NO}_x$  (refs 8, 9). A small contribution of  $0.015 \text{ g m}^{-2} \text{ yr}^{-1}$  was estimated from burning due to forest clearance<sup>10</sup> up to the year 1910. The source strength (total emissions divided by area of the contiguous US) is plotted as a line in Fig. 2a and agrees reasonably well with the observed changes. The location of England and Benelux about the North Sea does not allow such a simple calculation for the European data, but the display in Fig. 3 suggests the increase in nitrate deposition over the past 100 yr is smaller in western Europe than in eastern North America.

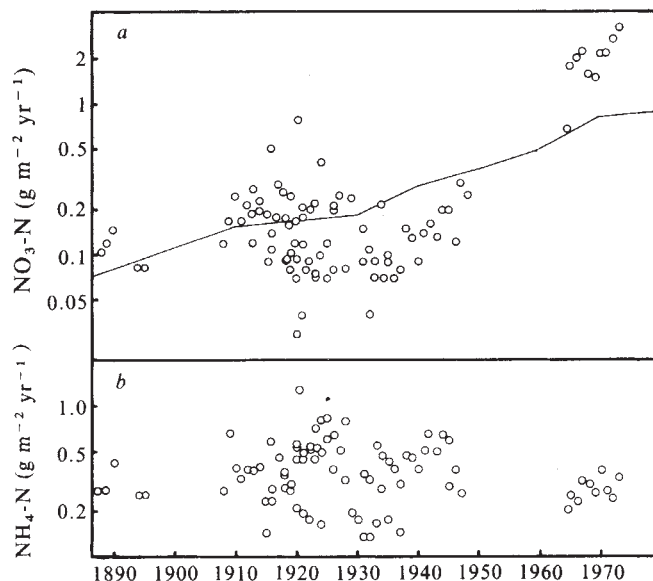


Fig. 2 Trends in the amount of nitrate (a) and ammonium (b) ion brought down annually by rainfall in eastern North America. All but the most recent data come from agricultural sources<sup>17–26</sup>. The line plotted through the nitrate ion data indicates the estimate of the source strength for nitrogen as nitrogen oxides<sup>27,28</sup>. There may have been a slight change in precipitation amount over the past 100 yr, but this would probably be a decrease and certainly  $<10\%$  (ref. 29).

These changes and the fact that there is substantial agreement between emissions of oxides of nitrogen and the deposition in the US reaffirm that anthropogenic combustion is largely responsible for the observed changes. The anthropogenic contributions are thus so large that natural contributions to the nitrate ion concentrations in rainfall are presently almost negligible. This is in agreement with recently proposed global cycles<sup>4,11</sup>, but is in conflict with the higher soil sources of  $\text{NO}_x$  adopted by earlier workers<sup>12,13</sup>. Charleson and Rodhe<sup>14</sup> have recently noted how the lack of knowledge of the  $\text{HNO}_3$  cycle makes our understanding of the natural pH of rainfall difficult, but these results suggest that the levels of nitrate ion in US rainfall have now become greater than those of the ammonium ion, so that it can no longer neutralize the atmospheric nitric

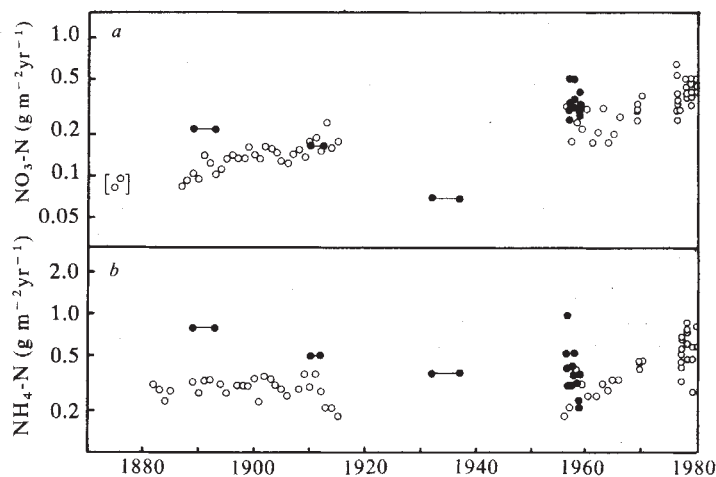


Fig. 3 Trends in the amount of nitrate (a) and ammonium (b) ion brought down annually in Belgium<sup>30,31</sup>, England<sup>2,5,15,30–34</sup>, and The Netherlands<sup>28,29</sup>. ●—●, Precipitation weighted annual means and the periods over which they are averaged. ○, English data; ●, continental data. The two bracketed points in the nitrate graph are displaced 20 yr from 1855 to 1856. The rainfall amount at Rothamsted shows no obvious long-term trend over the past 100 yr (ref. 1).



acid as it may have done in pre-industrial times. Figure 2a implies an increase in nitrate ion deposition over eastern North America of  $\sim 0.06$  equiv  $\text{m}^{-2} \text{yr}^{-1}$  since 1880. In regions receiving a 100-cm annual precipitation this would imply a precipitation weighted mean  $p_{\text{H}}$  shift from 5.6 to about 4.2 in carbon dioxide equilibrated water. Note that the present day deposition of sulphate ions ( $0.09$  equiv  $\text{m}^{-2} \text{yr}^{-1}$ ;  $1.2$  g (S)  $\text{m}^{-2} \text{yr}^{-1}$  wet deposition and  $0.25$  g (S)  $\text{m}^{-2} \text{yr}^{-1}$  dry deposition<sup>6</sup>) is not very much larger than the current deposit of nitrate ( $0.065$  equiv  $\text{m}^{-2} \text{yr}^{-1}$ ), reaffirming that the nitrate component of rain makes an important contribution to its acidity. Agricultural records do hint at an increase in sulphur deposition, but

problems experienced in early analysis make these changes less clear than those for the nitrate and ammonium ion. The relative contribution of the nitrate and sulphate components of western European precipitation are different. In the English east Midlands, for example<sup>15</sup>, the deposits are:  $\text{NO}_3^-$ ,  $0.03$ ;  $\text{SO}_4^{2-}$ ,  $0.055$  equiv  $\text{m}^{-2} \text{yr}^{-1}$ , showing a slightly smaller contribution from the nitrogen component than in the North American observations. This difference might be expected from the differing patterns of fuel usage, as the number of equivalents provided by the  $\text{NO}_x$  component of the UK fuel emission is less than one-quarter of that derived from sulphur, while in the US it is about one-half.<sup>16</sup>

Received 11 March; accepted 1 June 1982.

1. Brimblecombe, P. & Pitman, J. I. *Tellus* **32**, 261–267 (1980).
2. Russell, E. J. & Richards, E. H. *J. agric. Soc.* **9**, 309–337 (1919).
3. Liebig, J. *Chemistry and its Applications to Agriculture and Physiology* 3rd edn (Taylor and Walton, London, 1843).
4. Soderlund, R. & Svensson, B. H. *SCOPE* **7**, 27–73 (1976).
5. Miller, N. H. *J. agric. Sci.* **1**, 280–303 (1905).
6. Galloway, J. N. & Whelpdale, D. M. *Atmos. Envir.* **14**, 409–417 (1980).
7. Hammond, A. L., Metz, W. D. & Maugh, T. H. II *Energy and the Future* (American Association for the Advancement of Science, 1973).
8. Mulhaier, J. L. in *Solid Residential Fuels* (eds. Cooper, J. A. & Malek, D.) 164–187 (Oregon Graduate Centre, 1982).
9. Perkins, H. C. *Air Pollution* (McGraw-Hill, New York, 1974).
10. Wong, C. S. *Science* **200**, 197–200 (1978).
11. Stedman, D. H. & Shetter, R. E. *Adv. Envir. Sci. Technol.* (in the press).
12. Robinson, E. & Robbins, R. C. *Air Pollut. Control Ass. J.* **20**, 303 (1970).
13. Robinson, E. & Robbins, R. C. in (ed. Singer, S. F.) *The Changing Global Environment* 50–64 (Springer, New York, 1970).
14. Charlson, R. J. & Rodhe, H. *Nature* **295**, 683–685 (1982).
15. Martin, A. & Barber, F. R. *CEGB (MR) Rural Rain Sampling Sites: Monthly Mean Concentration Values 1975–1980* (CEGB MID/SSD/81/0046/M, 1981).

16. Bach, W. *Rev. Geophys. Space Phys.* **14**, 429–474 (1976).
17. Collinson, R. C. & Mensching, J. E. *New York St. agric. exp. Stat. tech. Bull.* No. 193 (1932).
18. Failyer, G. H. & Willard, J. T. *2nd Ann. Rep. Kansas agric. Stat.* 132–132 (1889).
19. Freeman, V. F. *J. Am. Soc. Agron.* **16**, 356–8 (1924).
20. Likens, G. E. *Biogeochemistry of a Forested Ecosystem* (Springer, Berlin, 1977).
21. Moores, C. A., McIntyre, W. H. & Young, J. B. *Univ. Tennessee agric. exp. Stat. Bull.* No. 138 (1927).
22. Shutt, F. T. *Proc. Trans. R. Soc. Can.* **4**, 55–59 (1910).
23. Shutt, F. T. *Proc. Trans. R. Soc. Can.* **8**, 83–87 (1914).
24. Shutt, F. T. & Dorrance, R. L. *Proc. Trans. R. Soc. Can.* **11**, 63–67, (1917).
25. Shutt, F. T. & Medley, B. *Proc. Trans. R. Soc. Can.* **19**, 1–10 (1925).
26. Wilson, B. D. *J. Am. Soc. Agron.* **18**, 1208–1212 (1926).
27. Cavender, J. H., Kircher, D. S. & Hoffman, A. J. *US EPA Publ.* No. AP-115 (1973).
28. Clark, T. L. *Atmos. Envir.* **14**, 960–961 (1980).
29. Wahl, E. W. & Lawson, T. L. *Mon. Weath. Rev.* **98**, 259–265 (1970).
30. *Tellus* **8**, 112 (1956).
31. Erikson, E. *Tellus* **4**, 215–232 (1952).
32. Jenkinson, D. S. *Rothamsted Exptl Stat Rept.* 1976, 103–109 (1977).
33. Dennington, V. N. & Chadwick, M. J. *J. appl. Ecol.* **15**, 303–315 (1978).
34. White, E., Starkey, R. S. & Saunders, M. J. *J. Appl. Ecol.* **8**, 743–749 (1971).

## Emissions of nitrous oxide from soils

J. M. Duxbury & D. R. Bouldin

Department of Agronomy, Cornell University, Ithaca, New York 14853, USA

R. E. Terry\* & R. L. Tate III†

University of Florida, AREC Belle Glade, Florida 33430, USA

Potential changes in the concentration of nitrous oxide ( $\text{N}_2\text{O}$ ) in the atmosphere have sparked considerable interest because of the proposed role of  $\text{N}_2\text{O}$  in regulating stratospheric ozone levels, and in contributing to the atmospheric greenhouse effect. A substantial portion of the atmospheric  $\text{N}_2\text{O}$  is thought to result from microbial transformations of inorganic forms of nitrogen in soils;  $\text{N}_2\text{O}$  is an intermediate in denitrification (reduction of  $\text{NO}_3^-$  to  $\text{N}_2$ ) and is formed during nitrification (oxidation of  $\text{NH}_4^+$  to  $\text{NO}_3^-$ ) in soils, although the mechanism is unclear. Several models have predicted that input of nitrogen into cropland, either from commercial fertilizers or N-fixing leguminous crops, could sufficiently increase emissions of  $\text{N}_2\text{O}$  from soils to deplete stratospheric ozone levels<sup>1–3</sup> and raise average world temperatures<sup>4</sup>. We report here  $\text{N}_2\text{O}$  emissions from mineral and organic soil sites in New York and from organic soil sites in the Florida Everglades Agricultural Area.

Only a few studies of emissions of  $\text{N}_2\text{O}$  from soils have been carried out over extended time periods in commercial agricultural conditions. In the US Hutchinson and Mosier<sup>5</sup> measured a loss of  $2.6$  kg  $\text{N}_2\text{O}-\text{N ha}^{-1}$  during corn growth in typical Colorado conditions and emissions of  $\text{N}_2\text{O}$  ranging from  $6$  to  $40$  kg  $\text{N ha}^{-1} \text{yr}^{-1}$  have been reported for heavily fertilized, irrigated, vegetable fields in California<sup>6,7</sup>. Ryden found that  $3.3$  kg  $\text{N}_2\text{O}-\text{N ha}^{-1} \text{yr}^{-1}$  were evolved from fertilized perennial ryegrass in the UK<sup>8</sup>. These data exceed the Council for Agricultural Science and Technology's (CAST) estimated average value for cropland of  $1$  kg  $\text{N}_2\text{O}-\text{N ha}^{-1} \text{yr}^{-1}$  (ref. 9) and indicate the

need for further assessment of emissions of  $\text{N}_2\text{O}$  from agricultural fields.

Measurement of  $\text{N}_2\text{O}$  flux from sites in New York and Florida was by a chamber method<sup>10</sup>. Sampling locations within a site were defined for extended periods of time by a metal rim (60 or 76 cm diameter) sunk 5 or 10 cm into the soil. When a flux measurement was to be made, a chamber was completed by sealing an insulated top to the rim with a wide rubber band cut from tire innertubes. Chamber volume was 28, 56, or 80 l and tops were removed immediately after each flux measurement. Gas samples (5 or 10 ml) were withdrawn from the chamber at zero time and 10-min intervals up to 30 min or at 30-min intervals up to 1 h, and their  $\text{N}_2\text{O}$  contents were determined by gas chromatographic analysis using a  $^{63}\text{Ni}$  electron capture detector<sup>10</sup>. Analytical precision was routinely 1% at the 1 p.p.m. level. Changes in concentration of  $\text{N}_2\text{O}$  greater than 10 p.p.b. in 1 h (3% of ambient level, corresponding to a flux of about  $0.5$  g  $\text{N ha}^{-1} \text{day}^{-1}$ ) were considered significant. Frequency of flux measurement was based on the magnitude of observed fluxes and on environmental and management variables. Measurements were usually made daily during high flux periods.

Daily flux values were the mean of four or five simultaneous flux measurements from chambers spaced randomly (5–10 m apart) over the site. The coefficient of variation (CV) for daily flux measurements ranged from 0 to 224%, with means in the range 60–94% for New York and 43–56% in Florida. No relationship between flux magnitude and CV was found. Fluxes from individual chambers showed similar trends with time but were sometimes out of phase during high flux periods. Variability amongst chambers was consistently reduced when fluxes were summed over important flux periods, allowing significant effects of treatment and/or crop on  $\text{N}_2\text{O}$  emissions to be demonstrated.

Mineral soil sites in New York were on silt loam soils. Alfalfa and field corn were grown following established practices for the northeastern US and an unmanaged timothy grass-mixed weed stand was included for comparative purposes. Only corn received fertilizer N;  $20$  kg  $\text{N ha}^{-1}$  was applied to the seed bed at planting, but the major nitrogen supply came either from manure applied and ploughed down before planting ( $C_M$ ) or from a sidedressing of commercial liquid fertilizer applied 40 days after planting ( $C_F$ ). Annual emissions of  $\text{N}_2\text{O}$  from

Present addresses: \*Department of Agronomy, Brigham Young University, Provo, Utah 84602, USA; †Department of Soils and Crops, The State University of New Jersey, New Brunswick 08903, USA.

**Table 1** Emissions of nitrous oxide from cultivated soils

| Crop                    | N fertilizer treatment* | Maximum observed                                        | Annual N <sub>2</sub> O emissions |                     |
|-------------------------|-------------------------|---------------------------------------------------------|-----------------------------------|---------------------|
|                         |                         | daily N <sub>2</sub> O flux<br>(kg N ha <sup>-1</sup> ) | May 1979 – May 1980               | May 1980 – May 1981 |
| Mineral soils           |                         |                                                         |                                   |                     |
| Alfalfa                 | 0                       | 0.63                                                    | 4.2                               | 2.3                 |
| Field corn A            | 130 (M)                 | 0.22                                                    | 2.4                               | 2.9                 |
|                         | 130 (F)                 | 0.28                                                    | —                                 | 2.2                 |
| Field corn B            | 130 (M)                 | 0.19                                                    | —                                 | 3.8                 |
|                         | 130 (F)                 | 0.36                                                    | 1.6                               | 2.9                 |
| Timothy-weeds           | 0                       | 0.23                                                    | 1.7                               | 0.9                 |
| Organic soils           |                         |                                                         |                                   |                     |
| Onions (NY)             | 170                     | 4.5                                                     | 85                                | 72                  |
| Sweet corn              | 170                     | 2.9                                                     | 76                                | 152                 |
| Sugarcane (FL)          | 0                       | 3.1                                                     | 48                                | 7                   |
| St Augustine-grass (FL) | 0                       | 4.6                                                     | 97                                | 16                  |
| None (fallow) (FL)      | 0                       | 4.5                                                     | 165                               | 59                  |

\* Field corn sites received 20 kg N ha<sup>-1</sup> added as NH<sub>4</sub>NO<sub>3</sub> in the seedbed and 112 kg N ha<sup>-1</sup> supplied by either (NH<sub>2</sub>)<sub>2</sub>CO–NH<sub>4</sub>NO<sub>3</sub> (1:1) fertilizer sidedressed 40 days after planting (F), or manure (45 tonnes ha<sup>-1</sup>) applied before planting (M). A single fertilizer treatment (M or F) was applied to each site in 1979–80. In 1980–81 each site was subdivided into 10-m wide strips which received either the M or F treatment on an alternating strip basis. Fertilizer containing NH<sub>4</sub>NO<sub>3</sub> was broadcast on the organic soil sites in New York before planting.

these sites over the 2 yr period from 1 May 1979 to 30 April 1981 ranged from 1.5 to 4.2 kg N ha<sup>-1</sup> (Table 1). Nitrous oxide evolution from the N-fixing alfalfa site was greater than from the corn sites in the first year of the study, but in the second year, emissions increased in the order alfalfa < C<sub>F</sub> < C<sub>M</sub>. In both years, more N<sub>2</sub>O was evolved from the C<sub>M</sub> than from the C<sub>F</sub> treatment and the least N<sub>2</sub>O was lost from the timothy-mixed weed pasture. The pattern of N<sub>2</sub>O emissions from the four sites varied considerably. High fluxes of N<sub>2</sub>O were always associated with wet soil conditions, which promote denitrification, but the converse was not always true. Nitrous oxide evolution from the alfalfa and timothy-mixed weed sites was greatest during early spring thaws; for example, 66 and 57% of the 1979–80 annual total from the alfalfa and timothy plots, respectively, was evolved between 20 March and 7 April 1980. The highest observed daily flux was 0.63 kg N ha<sup>-1</sup> from the alfalfa plot on 20 March 1980 and 27% of the 1979–80 annual total was evolved in the two days 19–20 March.

Peak emissions of N<sub>2</sub>O from the corn plots occurred following fertilization. Over the two study years, 62–78% of the annual loss of N<sub>2</sub>O from the manured sites occurred in the 9 weeks following manure application (end of April). Evolution of N<sub>2</sub>O from site B (Table 1) following sidedressing of liquid fertilizer nitrogen on 25 June in 1979 and 1980 is shown in Fig. 1. Nitrogen was applied to alternate rows and the impact of fertilization was compared using four pairs of chambers. The chambers in each pair were placed opposite each other in adjacent fertilized and unfertilized rows. In 1979, fluxes of N<sub>2</sub>O from fertilized rows were elevated compared with unfertilized rows for 9 weeks following fertilization. During this time, N<sub>2</sub>O emissions were 0.80 and 0.18 kg N ha<sup>-1</sup> for fertilized and unfertilized rows, respectively. A similar pattern was observed in 1980 but wetter soil conditions resulted in evolution of 3.40 and 0.35 kg N<sub>2</sub>O–N ha<sup>-1</sup> from fertilized and unfertilized rows, respectively, during the 7-week period that fluxes were elevated by fertilizer. These differences were significant at the 5% confidence level.

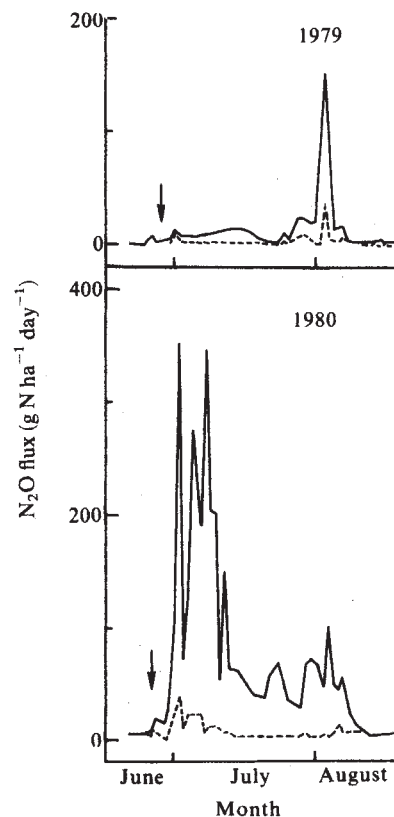
Significant effects of crop and/or treatment on N<sub>2</sub>O emissions could be demonstrated for selected time intervals within a year because, with the exception of the alfalfa and timothy plots, peak flux periods did not coincide. When cumulative fluxes from individual chambers were compared, N<sub>2</sub>O emissions from the alfalfa site were significantly greater (5% confidence level)

than from all other sites for the period 24 March to 10 April. Similar results were obtained for the corn (M) and corn (F) sites for the periods 1 May to 24 June and 25 June to 10 August, respectively. Statistical comparison of annual N<sub>2</sub>O emission totals could not be made because chambers were moved between peak flux periods.

Cultivated organic soil sites were included in the present study because most of the estimated 600–1,200 kg ha<sup>-1</sup> of nitrogen mineralized each year<sup>11,12</sup> appears to be denitrified. The study sites were located in a fallow field (bare) and fields cropped to St Augustine-grass (*Stenotaphrum secundatum* (Walt) Kuntz) and sugarcane (*Saccharum* spp.) in the Florida Everglades, and in fields cropped to onions (*Allium*) and sweet-corn (*Zea mays* L. Rugosa) in New York.

Both daily fluxes and annual emissions of N<sub>2</sub>O from the organic soil sites were considerably higher than from the mineral soil sites. Maximum daily fluxes of N<sub>2</sub>O ranged from 2 to 5 kg N ha<sup>-1</sup> and annual totals from 7 to 165 kg N ha<sup>-1</sup> (Table 1). High N<sub>2</sub>O fluxes were sustained for substantial periods of time (Fig. 2); for example, the daily flux of N<sub>2</sub>O from the fallow field in Florida was above 1 kg N ha<sup>-1</sup> for 27 days in July 1979. Peak emissions of N<sub>2</sub>O occurred during the wet summer months in Florida and the wet spring and fall periods in New York. A striking feature of the emission pattern in New York was daily fluxes commonly in the range of 0.2–0.4 kg N ha<sup>-1</sup> during January and February, despite the fact that the surface soil was generally frozen. Soil atmosphere measurements indicated that N<sub>2</sub>O was generated in the section of the soil profile that was unfrozen and above the water table.

Annual N<sub>2</sub>O emissions from the Florida sites differed substantially in the two study years (Table 1). The dramatic reduction in the second year was attributed to an unusually low amount of precipitation (92 cm compared with 157 cm in the first year and a 54 yr average of 145 cm). Annual emissions varied considerably with crop and showed the same relative pattern in both years increasing in the order sugarcane < grass <



**Fig. 1** Emissions of nitrous oxide from a corn field following sidedressing with (NH<sub>2</sub>)<sub>2</sub>CO–NH<sub>4</sub>NO<sub>3</sub> (1:1) at the rate of 112 kg N ha<sup>-1</sup>. Fertilizer was applied to alternate rows at the time indicated by the arrow. Solid and dashed lines are the N<sub>2</sub>O flux from the fertilized and unfertilized rows, respectively.



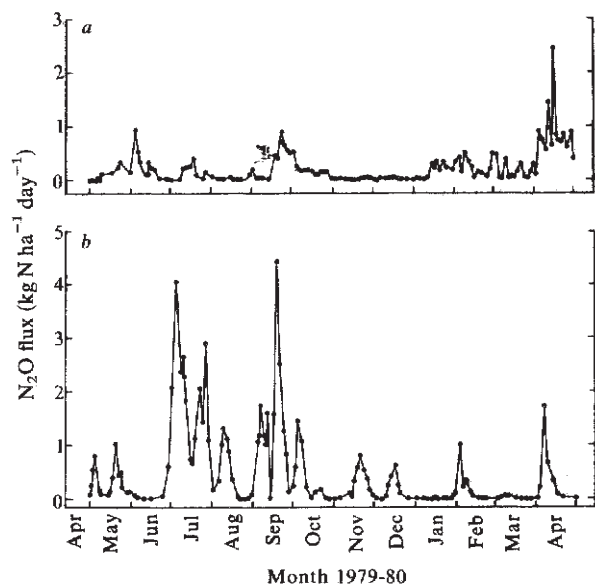


Fig. 2 Emissions of nitrous oxide from cultivated organic soils. a, Sweet corn field (New York); b, fallow field (Florida).

fallow. At the 5% confidence level, the differences between the sites were all significant in the first study year. In the second year,  $N_2O$  loss from the fallow field was significantly higher than from the other sites, which were not different. The large loss from the fallow field compared with the cropped fields was probably the combined result of wetter soil and higher soil nitrate levels in the absence of plants.

In contrast to the report of Ryden<sup>8</sup> none of the soils we studied were ever found to be sinks for atmospheric  $N_2O$ . Over the 2-yr period of the study ~13,000  $N_2O$  flux measurements were made. A likely explanation for this difference is that in our study soil nitrate levels never reached the very low level ( $\leq 1 \mu g N g^{-1}$  soil) that Ryden found to be associated with sink activity.

The crucial factor in assessing the impact of agriculture on emissions of  $N_2O$  from soils is the extent of perturbation from the undisturbed environment. True baseline data for present day agricultural lands probably cannot be obtained because man has, in the recent past, modified almost all of the land surface in the major agricultural regions of the world, and has increased precipitation inputs of nitrogen. Nevertheless, a survey of soils in Iowa that have no history of fertilizer N treatment showed that annual emissions of  $N_2O$  generally did not exceed  $0.5 kg N ha^{-1}$  (A. M. Blackmer and J. M. Bremner, personal communication). Our measurements at a forested mineral soil site in New York and at an organic soil site in a Florida Everglades Conservation Area gave emissions of about 0.9 and  $1 kg N_2O-N ha^{-1}$ , respectively, for the calendar year 1980. Compared with these values, it can be concluded that agriculture caused  $N_2O$  emissions from soils to increase up to fourfold for mineral soil sites in New York and by as much as two orders of magnitude from the organic soil sites studied.

It is difficult to place the observations on  $N_2O$  emission from agricultural lands in a global perspective because the data base is limited to a few agricultural systems and temperate climates<sup>5-8,13-19</sup>. Nevertheless, from an evaluation of the present data base, we conclude that forests and grasslands, which lose the least  $N_2O$  on a unit area basis, may well be the major contributors to the global total because of their large areas. Conversely, the contribution of vegetable cropland and drained organic soils, which lose high amounts of  $N_2O$ , is probably small because of their small areas. It is clear, however, that there is a need for further research on  $N_2O$  emissions from soils to focus on unmanaged lands and the least intensively managed agricultural lands, which has not been the trend to date.

We thank A. Bekerie, S. Voghel, D. Manning, S. LaPointe, D. Miller, W. Reeve and the late G. Dresser for technical

assistance. This research was supported in part by a grant from the NSF.

Received 1 March; accepted 19 May 1982.

1. Crutzen, P. J. & Ehhalt, D. H. *Ambio* **6**, 112-117 (1977).
2. Liu, S. C., Cicerone, R. J., Donahue, T. M. & W. L. Chameides, *Tellus* **29**, 251-263 (1977).
3. McElroy, M. B., Wofsy, S. C. & Yung, Y. L. *Phil. Trans. R. Soc. B* **277**, 159-181 (1977).
4. Yung, Y. L., Wang, W. C. & Lacis, A. A. *Geophys. Res. Lett.* **3**, 619-621 (1976).
5. Hutchinson, G. L. & Mosier, A. R. *Science* **205**, 1125-1127 (1979).
6. Ryden, J. C., Lund, L. J., Letey, J. & Focht, D. D. *Soil Sci. Soc. Am. J.* **43**, 110-118 (1979).
7. Ryden, J. C. & Lund, L. J. *Soil Sci. Soc. Am. J.* **44**, 505 (1980); *J. envir. Qual.* **9**, 387-393 (1980).
8. Ryden, J. C. *Nature* **292**, 235-237 (1981).
9. Council for Agricultural Science and Technology *Effect of Increased Nitrogen Fixation on Stratospheric Ozone* (Rep. 53, Iowa State University, Ames, 1976).
10. Terry, R. E., Tate, R. L. III & Duxbury, J. M. *J. Air Pollut. Contr. Ass.* **31**, 1173-1176 (1981).
11. Guthrie, T. F. & Duxbury, J. M. *Soil Sci. Soc. Am. J.* **42**, 908-912 (1978).
12. Terry, R. E. *Soil Sci. Soc. Am. J.* **44**, 747-750 (1980).
13. Burford, J. R. & Hall, K. C. *Argic. Res. Council, Letcombe Lab. Ann. Rep.* **85** (1976).
14. Rolston, D. E., Fried, M. & Goldammer, D. A. *Soil Sci. Soc. Am. J.* **40**, 259-266 (1976).
15. McKenney, D. J., Wade, D. L. & Findlay, W. I. *Geophys. Res. Lett.* **5**, 777-780 (1978).
16. Rolston, D. E., Hoffman, D. L. & Toy, D. W. *Soil Sci. Soc. Am. J.* **42**, 863-869 (1978).
17. Denmead, O. T., Freney, J. R. & Simpson, J. R. *Soil Sci. Soc. Am. J.* **43**, 726-728 (1979).
18. Bremner, J. M., Robbins, S. G. & Blackmer, A. M. *Geophys. Res. Lett.* **7**, 641-644 (1980).
19. McKenney, D. J., Shuttleworth, K. F. & Findlay, W. I. *Can. J. Soil Sci.* **60**, 429-438 (1980).

## A predatory slime mould

David R. Waddell\*

Department of Biology, Princeton University, Princeton, New Jersey 08544, USA

I report here the discovery of a single isolate of a new species of cellular slime mould, *Dictyostelium caveatum*, in bat guano from Blanchard Springs Cavern, Arkansas. Amoebae of this species form common aggregates with all species that have been tested and interfere with their morphogenesis in a dramatic way. The *D. caveatum* amoebae induce a delay of morphogenesis while they consume all of the amoebae of the other species by phagocytosis. Morphogenesis is reinitiated and fruiting bodies containing only *D. caveatum* spores are erected.

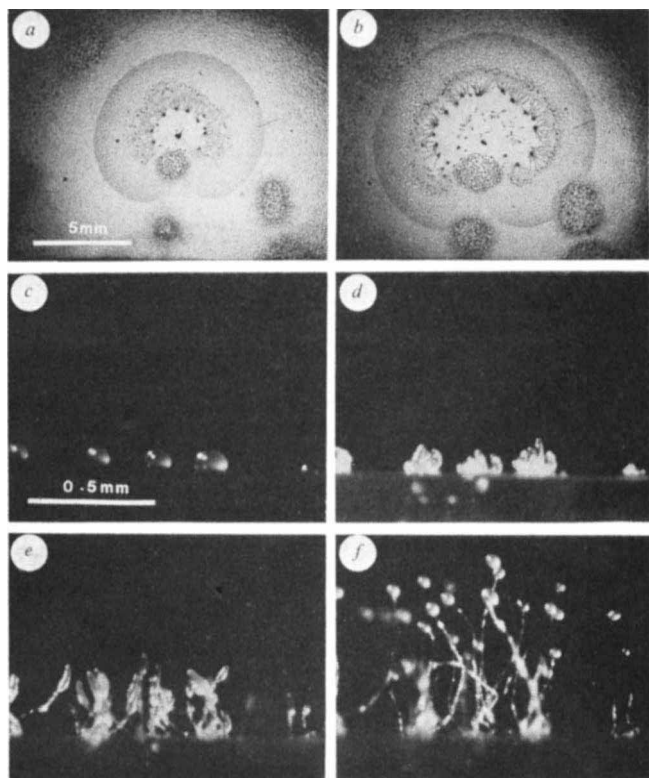
The bat guano deposit from which *D. caveatum* was isolated is situated in darkness where the temperature is a constant  $15^\circ C$  and the relative humidity is always 100%. Seven other species were isolated from the same deposit. A single isolate of *D. caveatum* was obtained which represented less than 1% of the total clones recovered. Stages in the life cycle are shown in Fig. 1.

The amoebae of this species respond in chemotactic tests<sup>1</sup> to *Polysphondylium violaceum* acrasin but do not respond to cyclic AMP, suggesting a close relationship to *Polysphondylium*. The presence of polar spore granules<sup>2</sup> and the high G+C content (30.6%) of the main band DNA<sup>3</sup> also suggest a close affinity to the *Polysphondylium*-related group defined by Traub and Hohl<sup>2</sup>. This group includes several *Dictyostelium* species that are classified in this genus because they lack whorls on their fruiting bodies. *D. caveatum* amoebae migrate as single cells without forming chains and streams (Fig. 1a, b). However, towards the end of aggregation, short radial streams can sometimes be observed.

*D. caveatum* amoebae follow the normal cellular slime mould life cycle<sup>4</sup> when growing on bacteria. However, when *D. caveatum* amoebae are washed free of bacteria and mixed with amoebae of other species, only *D. caveatum* spores can be recovered at the end of morphogenesis. This occurs in mixtures of *D. caveatum* with *Dictyostelium discoideum*, *Dictyostelium purpureum*, *Dictyostelium mucoroides*, *Dictyostelium polycephalum*, *Dictyostelium rosarium*, *Dictyostelium minutum*, *Polysphondylium pallidum*, *Polysphondylium violaceum*, *Acytostelium leptosomum* and an unidentified isolate in this laboratory designated LOL-1.

Present address: Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG.





**Fig. 1** Stages in the life cycle of *D. caveatum*. *a*, A low density suspension of *D. caveatum* and *P. pallidum* amoebae was spread on lactose-peptone agar ( $2 \text{ g l}^{-1}$  Bactopeptone (Difco),  $2 \text{ g l}^{-1}$  lactose,  $15 \text{ g l}^{-1}$  agar,  $0.272 \text{ g l}^{-1} \text{ KH}_2\text{PO}_4$ ,  $0.284 \text{ g l}^{-1} \text{ Na}_2\text{HPO}_4$ ) with *Escherichia coli* B/r. After 15–20 generations, each cell had formed a colony which was visible as a clearing in the bacterial lawn. Cells far enough from the feeding front of the colony starved and began to aggregate. *P. pallidum* amoebae formed extensive streams as they migrated to aggregation centres. The large colony with the discrete feeding edge in the centre is a *P. pallidum* clone. In contrast, *D. caveatum* amoebae migrate to centres as single cells. A *D. caveatum* colony lies within the feeding edge of the *P. pallidum* clone. Note the curving in of the *P. pallidum* feeding edge near the *D. caveatum* clone. The diffuse feeding edge and relatively slower growth of the *D. caveatum* clones are characteristic of this species. The photograph in *a* was taken  $\sim 72 \text{ h}$  after the cultures were initiated. *b*, The same colonies  $8.5 \text{ h}$  later. A third round of aggregation was proceeding in the *P. pallidum* colony. The central aggregate and two of the aggregates that were forming  $8.5 \text{ h}$  earlier were arrested at the tight aggregate stage. The arrested aggregates eventually formed *D. caveatum* fruiting bodies. The other aggregates that had formed at the same time had proceeded to the early culmination stage of fruiting body formation. *c–f*, Different stages in the morphogenesis of a pure culture of *D. caveatum*. *D. caveatum* amoebae were collected from bacterial cultures before any clearing of the bacterial carpet was visible, washed free of bacteria by differential centrifugation and plated on non-nutrient agar ( $0.6 \text{ g l}^{-1} \text{ NaCl}$ ,  $0.75 \text{ g l}^{-1} \text{ KCl}$ ,  $0.3 \text{ g l}^{-1} \text{ CaCl}_2$ ,  $20 \text{ g l}^{-1}$  Bactoagar, Difco). Approximately 5–6 h later, aggregation began. By  $11.5 \text{ h}$  (*c*), tight aggregates were formed. Tips began to appear at  $18.5 \text{ h}$  and were well developed by  $20 \text{ h}$  (*d*). Several tips commonly arose from a single aggregate. Culmination proceeded rapidly. By  $22.5 \text{ h}$  (*e*), stalks had risen to a height of  $0.4 \text{ mm}$  and by  $25 \text{ h}$  culmination was complete (*f*).

Clonal analysis of developing cultures initiated with *P. pallidum* and *D. caveatum* in a 99:1 ratio (Fig. 2) has established that conversion occurred during the aggregation stages and was completed by 35 h. The number of *D. caveatum* plaques recovered from the mixed culture increased exponentially after aggregation and before conversion was completed. The doubling times of the exponential phase in the two experiments presented in Fig. 2 were 4.8 and 6.4 h. For every *D. caveatum* cell that appeared two to three *P. pallidum* cells were lost.

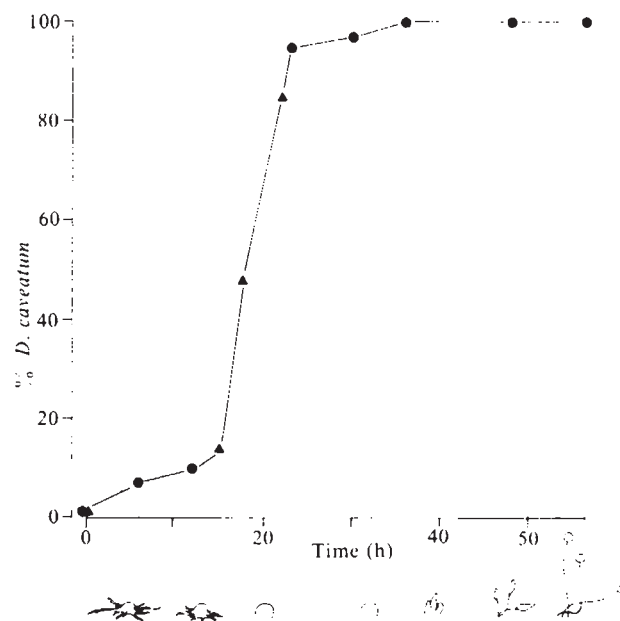
The efficiency and rapid kinetics of the conversion suggested that it might involve an infective particle similar to  $\kappa$  particles in *Paramecium*<sup>5</sup>. If this is the case, then the nuclear DNA of the converted cells should remain intact. To determine this, *D. caveatum* was mixed with *D. discoideum*, a species which has nuclear DNA of a different buoyant density from *D. caveatum*. The nuclear DNA of *D. discoideum* separated

into two bands on caesium chloride density gradients, a main band with a density of  $1.683 \text{ g cm}^{-3}$  and a satellite with a density of  $1.690 \text{ g cm}^{-3}$  (Fig. 2). *D. caveatum* DNA forms a single band at  $1.690 \text{ g cm}^{-3}$ . When the two species were mixed at an initial ratio of one *D. caveatum* cell per 100 *D. discoideum* cells and the nuclear DNA isolated after 12 h, two bands were observed. However, after 23 h only DNA which bands at 1.690 was detectable. This suggests that conversion involved the breakdown of the nuclear DNA of the prey species and rules out the possibility that only the passage of an infective particle could be responsible for conversion.

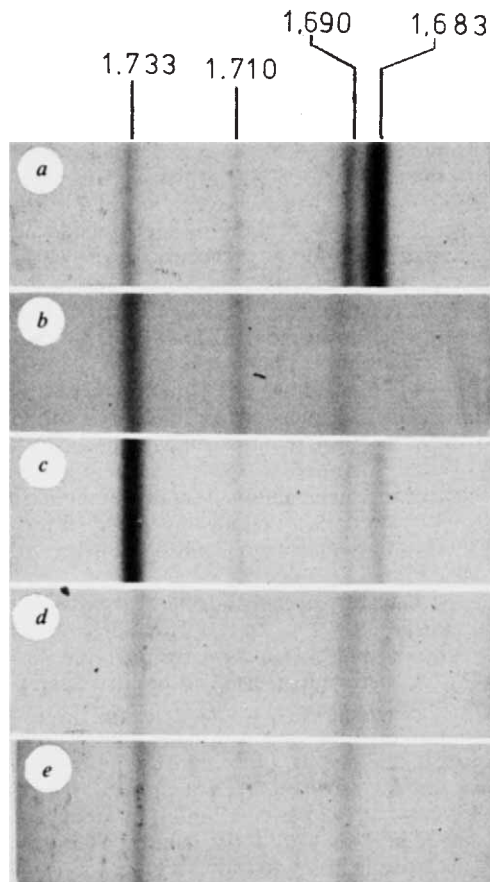
The most plausible alternatives are that: (1) *D. caveatum* cells kill and lyse prey amoebae and feed on the released nutrients; (2) *D. caveatum* amoebae fuse with the prey amoebae and in some way break down their nuclear DNA; or (3) the *D. caveatum* amoebae phagocytose the prey amoebae and digest them. When thin sections of aggregates fixed during conversion were examined in the electron microscope (Fig. 3*b*), large amoebae which contained smaller amoebae in phagosomes were observed. No cell fusions were observed. This suggests that the primary mechanism for conversion is phagocytosis of the prey amoebae.

A stage-specific delay of morphogenesis was associated with the conversion process (Fig. 1*a, b*). Three *P. pallidum* aggregates that were within the feeding front of the *D. caveatum* plaque in Fig. 1*a* were still arrested at the tight aggregate stage in Fig. 1*b*, although other *P. pallidum* aggregates that were formed at the same time or after, had progressed to the culminating stages of fruiting body construction. The arrested aggregates eventually formed *D. caveatum* fruiting bodies.

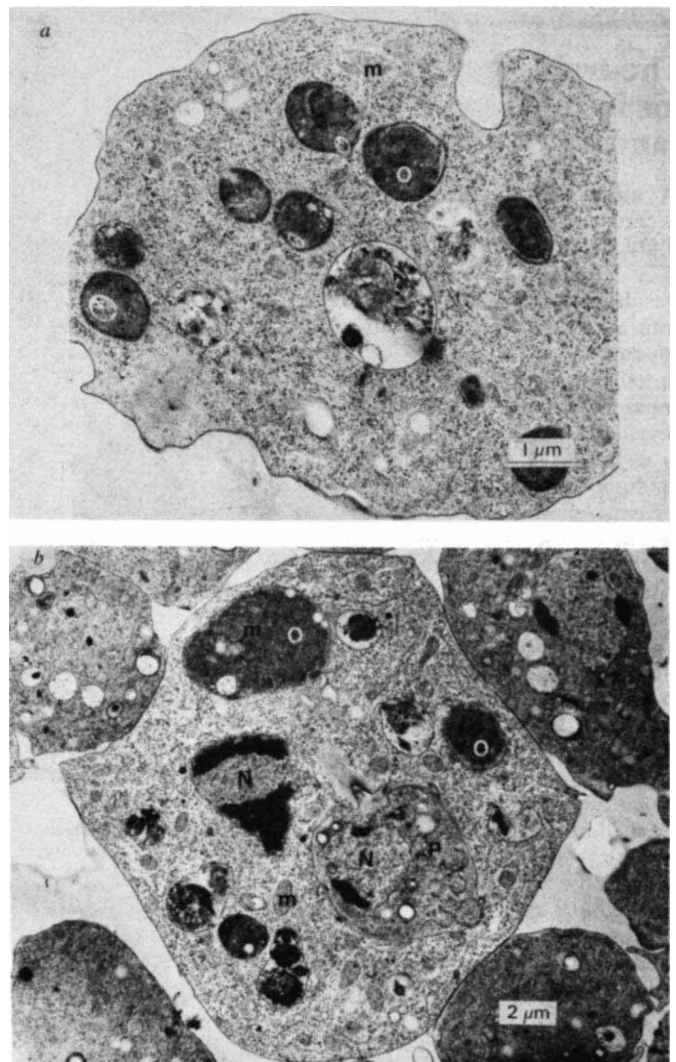
In experiments in which the initial ratio of prey to *D. caveatum* amoebae was varied, the length of the delay before morphogenesis progressed was inversely proportional to the initial percentage of *D. caveatum* cells. The time of tip formation corresponded closely to the time that conversion was complete by clonal analysis. However, when *D. caveatum* was



**Fig. 2** Kinetics of conversion of *P. pallidum* by *D. caveatum*. Amoebae of both species were collected from separate cultures growing in association with *E. coli* B/r. The amoebae were washed free of bacteria and counted. The two species were mixed in a 99:1 ratio of *P. pallidum*:*D. caveatum*.  $1.5 \times 10^7$  cells of this mixture were spread on each of a series of non-nutrient agar plates (Corning 35 mm, 25,000). At each time point the cells from two of the plates were collected separately and made up to a total volume of 10 ml. This suspension was diluted  $\times 4,000$ ,  $\times 40,000$  and  $\times 400,000$  and 0.4-ml of the dilutions was spread on lactose-peptone agar plates ( $100 \times 15 \text{ mm}$ ) with *E. coli* B/r. The colony morphology of the two species are distinct (Fig. 1). The ratio of the two types of colonies was determined for the duplicate set of plates. The results of two separate experiments are presented (●, ▲).



**Fig. 3** Caesium chloride density centrifugation of nuclear DNA isolated from *D. discoideum*, *D. caveatum* and at different times during the conversion of a *D. discoideum* culture by *D. caveatum*. Cultures of *D. discoideum* and *D. caveatum* were grown in association with *E. coli* B/r. Nuclear DNA was isolated following the procedure of Thomas *et al.*<sup>6</sup>. For the conversion experiment, vegetative amoebae of both species were washed free of bacteria and a 99% : 1% mixture was dispersed as before. After 12 and 23 h, the developing cultures were collected, the nuclei isolated and the DNA purified. *a*, *D. discoideum* nuclear DNA. Main band Density ( $\rho$  g cm<sup>-3</sup>) = 1.683, satellite  $\rho$  = 1.690. *M. lysodeikticus* DNA ( $\rho$  = 1.733) was added as a marker. The band at 1.710 is *E. coli* DNA which was the food organism for the slime moulds. *b*, *D. caveatum* nuclear DNA. Main band  $\rho$  = 1.690. *c*, Mixture of DNA from *D. caveatum* and *D. discoideum*. *d*, DNA isolated from cells lysed 12 h after initiating morphogenesis with a 99 : 1 mixture of *D. discoideum* : *D. caveatum*. *e*, DNA isolated from cells lysed 23 h after initiating morphogenesis with a 99 : 1 mixture of *D. discoideum* and *D. caveatum*.



**Fig. 4** *a*, Electron micrograph of *D. caveatum* vegetative cell. *D. caveatum* amoebae grown in association with *E. coli* B/r, were washed free of bacteria and allowed to permeate a Millipore filter (8 µm pore size) for 1 h. The cells were fixed and prepared<sup>7</sup> for electron microscopy. The sections were stained with uranyl acetate and lead citrate. *b*, Electron micrograph of a *D. caveatum* amoeba with an ingested *P. pallidum* amoeba. *P. pallidum* and *D. caveatum* amoebae grown in association with *E. coli* were washed free of bacteria and combined in a 99 : 1 ratio and dispersed on non-nutrient agar. After 12 h individual aggregates were cut away along with a small piece of agar and fixed in glutaraldehyde and osmium tetroxide and treated as above for electron microscopy. N, nucleus; P, phagocytosed amoeba; m, mitochondria; O, lysosomal-like organelle.

mixed with *D. discoideum* and *D. purpureum* at very low ratios (1:10,000), morphogenesis progressed past the tight aggregate stage. In both species migrating slugs which initially exhibited typical *D. discoideum* and *D. purpureum* behaviour were formed. *D. purpureum* slugs were formed within 6.5 h after the amoebae were dispersed on non-nutrient agar. After migrating for about 6 h the slugs stopped producing stalk and then rounded up, and lost their tips. *D. discoideum* slugs were not formed until 17.5 h after dispersal of the amoebae. The *D. discoideum* slugs migrated for 12 h before stopping, rounding up and losing their tips. Subsequently, the arrested mounds of both species reformed tips and constructed *D. caveatum* fruiting bodies ~60 h after the cultures were initiated. Although the two species were mixed with *D. caveatum* at the same initial ratio; the time that morphogenesis became visibly aberrant was very different; *D. purpureum* morphogenesis was arrested after 12.5 h but *D. discoideum* morphogenesis was not arrested until 29.5 h.

As only three to six doublings of *D. caveatum* could have occurred by the time morphogenesis became aberrant, the majority of the cells in the slugs were still the prey species. This suggests that *D. caveatum* amoebae use an inhibitory

mechanism that arrests differentiation or morphogenesis in the prey amoebae before phagocytosis.

I thank John Bonner, Edward Cox, Kei Inouye and Donna Bozzone for help, James Ramey and the USDA for samples of bat guano, Barbara Bamman, Hannah Suthers and Kate Schenk for technical help and Tova Rothschild for suggesting the name for the species. D.R.W. was supported by a NIH postdoctoral fellowship (5-T32-CA009167-05). The work was supported by NIH research grant GM 17856-09 to John Bonner.

Received 22 January; accepted 19 May 1982.

1. Pan, P., Hall, E. M. & Bonner, J. T. *Nature new Biol.* **237**, 181 (1972).
2. Traub, F. & Hohl, H. R. *Am. J. Bot.* **63**, 664 (1974).
3. Dutta, S. K. & Mandel, M. J. *Protozool.* **19**, 538 (1972).
4. Bonner, J. T. *The Cellular Slime Molds* (Princeton University Press, 1967).
5. Soldo, A. T. in *Paramecium: A Current Survey* (ed. Van Wagtenonk) 377-442 (Elsevier, Amsterdam, 1974).
6. Thomas, C. A., Berns, K. & Kelly, J. in *Procedures in Nucleic Acid Research* Vol. 1 (eds Cantoni, G. & Davies, D.) 535-537 (Harper & Row, New York, 1966).
7. Swanson, J., Taylor, D. L. & Bonner, J. T. *J. ultrastruct. Res.* **75**, 243 (1981).



## The trained circling rat: a model for inducing unilateral caudate dopamine metabolism

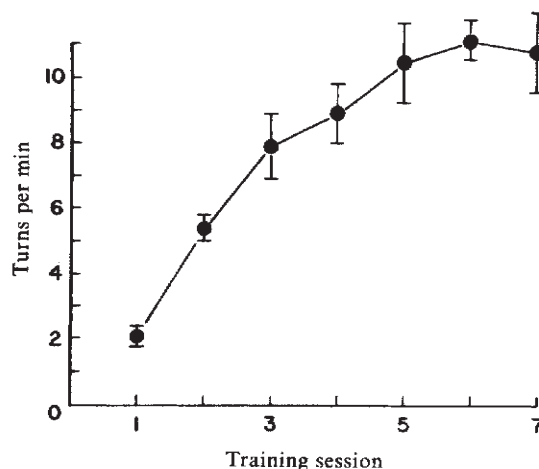
Bryan K. Yamamoto & Curt R. Freed

Departments of Medicine and Pharmacology, Box C237, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

Despite considerable knowledge about brain dopamine in drug-induced and pathological states, the dynamic relationship of dopamine to voluntary motor behaviour is not established. In an effort to produce selective, lateralized changes in brain dopamine metabolism associated with movement, we have now developed an animal model in which normal rats were trained to circle for a sucrose/water reward. Turning direction was randomly assigned. The effect of circling on dopamine metabolism was studied by killing animals during the reinforcement period and measuring caudate dopamine and dihydroxyphenylacetic acid (DOPAC) concentrations in both sides of brain. Prior to turning, caudate dopamine and DOPAC levels were the same bilaterally. By contrast, animals killed after 20 min of circling showed a 67% increase in dopamine and a 46% increase in DOPAC concentrations in the caudate contralateral to the turning direction while the ipsilateral caudate showed no concentration changes. Such a rapid, massive unilateral increase in caudate dopamine metabolism suggests that dopamine neurones in contralateral caudate are selectively activated during circling behaviour. This model should be useful for further study of the relationships between neurotransmitter activity, regional brain physiology and voluntary movement, without the need for drugs or brain lesions.

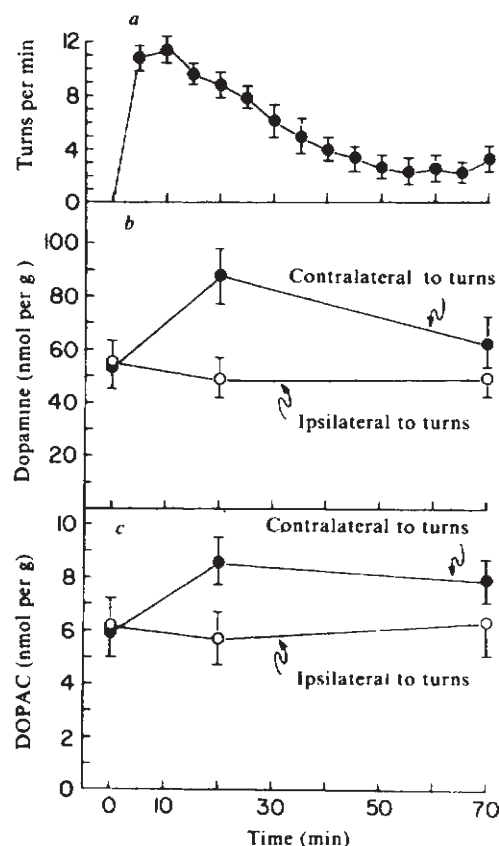
Circling behaviour is an extensively studied area of the neurosciences and has recently been reviewed<sup>1</sup>. Studies have used lesions, stimulation of various brain areas and pharmacological manipulations in an effort to understand the neural basis of motor behaviour. Most workers have focused on the dopamine-containing nigrostriatal system. Ungerstedt showed that animals with unilateral 6-hydroxydopamine lesions of the substantia nigra would circle in the direction of the lesioned side after amphetamine administration, presumably because of enhanced dopamine release in the caudate on the intact side<sup>2</sup>. In this lesion model, a receptor agonist such as apomorphine will produce turns in the direction opposite to the lesions, possibly because of the development of denervation supersensitivity on the lesioned side. Glick and co-workers have noted that some rats preferentially turn to one side spontaneously and that this turning behaviour can be increased by amphetamine<sup>3,4</sup>. Following amphetamine, these animals show an increase in dopamine in the caudate contralateral to the turning direction. Results from these and other experiments have suggested that striatal dopamine has a role in circling behaviour. Because all of these paradigms have used drugs or lesions or animals with natural behavioural asymmetries, these previous studies do not directly address the relationship between dopamine metabolism and movement in the normal animal. We have therefore developed a behavioural model to study caudate dopamine metabolism in animals trained to turn in an arbitrary direction without the use of drugs or lesions.

Eighteen male Sprague-Dawley rats weighing 250–350 g were water deprived for 24 h before training. All rats were randomly assigned to be trained to turn 360° in the left-hand or right-hand direction. The training apparatus was a 24-cm diameter circular drum with 25-cm vertical walls. The drum was equipped with a liquid dispenser (Lafayette) calibrated to deliver 0.05 ml of a 10% w/w sucrose/water per reinforcement. Following water deprivation, rats were placed in the drum and sucrose solution was released from the dispenser. All rats quickly located the fluid. The training protocol consisted of rewarding each rat by successive approximations to a full turn in the prescribed direction. Initially, a quarter turn in the



**Fig. 1** Acquisition of trained turning behaviour in six animals. Water-deprived animals were trained daily using a sucrose/water reward. Animal performance plateaued after 7 days of reinforcement training.

appropriate direction was reinforced. When this response was acquired, a 180° turn was then necessary for reward. Subsequently, 270° and 360° turns were required. Rats learned this procedure within 20–30 min and the first training session continued until each animal had performed 100 turns. Training was conducted daily for 7 days according to a continuous



**Fig. 2** Turning behaviour and caudate dopamine metabolism in trained rats. Trained animals were placed in the reinforcement apparatus at time 0. *a*, Turning began rapidly after water reinforcement was started. Data are mean  $\pm$  s.e.m. for six animals. *b*, *c*, Caudate dopamine and DOPAC concentrations were measured by killing six animals at each time point. Values are mean  $\pm$  s.e.m. Contralateral dopamine and DOPAC concentrations are significantly elevated compared with ipsilateral concentrations at 20 min ( $P < 0.01$ , both values) and 70 min ( $P < 0.05$ , both values) after starting circling. Concentrations are in nanomol per g wet weight caudate. Statistics were determined by Student's paired *t*-test.



reinforcement schedule. Figure 1 shows the acquisition curves for this behaviour. One hundred turns were required during the first 3 days of training while 150 complete turns were required for the subsequent 4 days of training. During this time, the only water provided was that received during the training session. *Ad libitum* access to Purina Laboratory Chow was available in the home cage. This deprivation schedule enabled all rats to maintain at least 85% of their original body weight.

We found that after turning behaviour was acquired, the animals retained the response for at least 1 month without regular training sessions. Minimal training was required to restore the previous maximum turning intensity. After animals had learned the turning paradigm, they were killed at different times during circling behaviour and caudate dopamine and DOPAC measured. Rats were killed by decapitation, and left and right caudate were dissected and assayed for dopamine and DOPAC by HPLC with electrochemical detection, using trichloroacetic acid as the ion-pairing agent according to the method of Freed and Asmus<sup>5</sup>.

The relationship between turning and dopamine metabolism is presented in Fig. 2. Figure 2a shows the time course of turning for animals permitted to follow the full turning protocol. No circling behaviour occurred during the baseline period before water reinforcement was made available. When the animals were transferred to the reinforcement apparatus, turning began immediately with high intensity. Turning rate declined after ~30 minutes and approached a low frequency plateau by 70 min. Caudate nuclei contralateral and ipsilateral to the trained circling direction were analysed for dopamine and its major metabolite, DOPAC (see Fig. 2b, c). Dopamine and DOPAC concentrations significantly increased in the caudate contralateral to the trained turning direction at the time of peak turning intensity. Contralateral concentrations were increased both in comparison with baseline values and with simultaneous values from the ipsilateral side. After 70 min of reinforcement, circling had slowed and contralateral dopamine and DOPAC concentrations declined but were still significantly greater than ipsilateral concentrations.

Our findings of unilateral increases in both dopamine and DOPAC concentrations imply that both synthesis and turnover of caudate dopamine are selectively enhanced during voluntary turning. Previous studies of dopamine metabolism in brain of resting animals have suggested that there is a small, readily releasable high turnover pool of dopamine in the caudate and a large slow turnover storage pool<sup>6</sup>. Our results indicate that the trained circling rat is capable of unilaterally increasing total caudate dopamine concentration by 67% in 20 min. Although we have not localized this dopamine, such a massive increase suggests that the entire dopamine pool is involved. Because contralateral caudate DOPAC concentration is also elevated over the time course of circling, increased dopamine synthesis is accompanied by increased transmitter turnover. Dopamine concentration rose more rapidly than DOPAC concentration (67% increase in dopamine versus 46% for DOPAC at 20 min), but at 70 min, as turning subsided, dopamine concentration had fallen to only 20% greater than baseline while DOPAC remained 35% increased. As DOPAC is a metabolite of dopamine, this delayed fall-off of DOPAC concentration is expected.

We have obtained further support for asymmetric regulation of caudate dopamine metabolism in independent experiments. Using the trained rat model and the technique of *in vivo* electrochemistry, we have found that contralateral release of dopamine is relatively increased during circling<sup>7</sup>. Furthermore, we have assayed tyrosine hydroxylase activity at the same time points reported here using the method of Nagatsu<sup>8</sup> at a sub-saturating concentration of pterin cofactor (0.4 mM 2-amino-4-hydroxy-6-methyltetrahydropterin). Baseline activity was not significantly different in the two sides of brain (ratio of activity contralateral/ipsilateral =  $0.84 \pm 0.08$ ,  $n = 5$ ). After 20 min circling, contralateral caudate tyrosine hydroxylase activity increased 106% relative to the ipsilateral side (activity ratio

$1.73 \pm 0.21$ ,  $n = 5$ ,  $P < 0.01$  compared with baseline ratio). These data are part of a complete study of the kinetics of tyrosine hydroxylase activation in our circling rat model (M.E. Morgan, B.K.Y. and C.R.F., in preparation).

Ours is the first observation of lateralization of brain neurotransmitter metabolism during voluntary motor behaviour. Nieoullon *et al.* demonstrated enhanced dopamine release in caudate ipsilateral to unilateral sensory stimulation in anaesthetized cats<sup>9</sup>. In addition, their experiments showed a reduction in dopamine release from the contralateral caudate. Glick and co-workers showed relative increases in contralateral caudate dopamine concentration after amphetamine administration to unlesioned animals which showed side preferences<sup>3</sup>, but our experiments are the first evidence that such changes occur in normal animals executing an arbitrarily assigned behaviour. Furthermore, the time course of the changes in dopamine concentration parallel the motor behaviour. Our experiments do not address the question of the causal role of dopamine in movement. Although the maximum increase in contralateral dopamine concentration is observed at the time of maximum circling intensity and is reduced as turning subsequently slows, we cannot say whether dopamine has caused or is simply the result of the turning behaviour.

The traditional substantia nigra-lesioned rotating rat model has been useful for investigating nigrostriatal function and as a screening device for potential anti-parkinsonian drugs. However, the trained rat model has an important advantage for studying neurotransmitters in the central nervous system. It provides a method for exploring the relationships between behaviour and dopamine metabolism in the normal, undrugged and unlesioned brain. In addition, the trained turning rat may be useful for studying laterality in other brain areas linked to motor function and to other neurotransmitters such as  $\gamma$ -aminobutyric acid and acetylcholine.

Received 21 January; accepted 21 May 1982.

1. Pycock, C. J. *Neuroscience* **5**, 461–514 (1980).
2. Ungerstedt, U. *Acta physiol. scand. Suppl.* **367**, 69–93 (1971).
3. Glick, S. D., Jerussi, T. P. & Zimmerman, B. in *Lateralization in the Nervous System* (ed. Harnad, S.) 213–249 (Academic, New York, 1977).
4. Jerussi, T. P. & Glick, S. D. *Psychopharmacology* **47**, 249–260 (1976).
5. Freed, C. R. & Asmus, P. A. *J. Neurochem.* **32**, 163–168 (1979).
6. Javoy, F. & Glowinski, J. *J. Neurochem.* **18**, 1305–1311 (1971).
7. Yamamoto, B. K., Lane, R. F. & Freed, C. R. *Life Sci.* **30**, 2155–2162 (1982).
8. Nagatsu, T., Levitt, M. & Udenfriend, S. *Analyt. Biochem.* **9**, 122–126 (1964).
9. Nieoullon, A., Cheramy, A. & Glowinski, J. *Nature* **266**, 375 (1977).

## Brain grafts reverse hypogonadism of gonadotropin releasing hormone deficiency

D. T. Krieger\*, M. J. Perlow\*§, M. J. Gibson\*, T. F. Davies\*, E. A. Zimmerman†, M. Ferin† & H. M. Charlton‡

\* Departments of Medicine and Neurology, Mount Sinai Medical Center, New York, New York 10029, USA

† Departments of Neurology and Obstetrics and Gynecology, College of Physicians & Surgeons, New York, New York 10032, USA

‡ Department of Human Anatomy, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK

**Hypogonadism in the mutant *hpg* mouse is characterized by a deficiency of hypothalamic gonadotropin releasing hormone (GnRH)<sup>1</sup>. Affected male mice exhibit immature reproductive organs, small abdominal testes and low pituitary and plasma gonadotropin concentrations. Recent studies have demonstrated the potential of fetal brain transplants to establish functional connections with host tissues<sup>2</sup>. We therefore sought to use this approach to correct the *hpg* deficit. Fetal preoptic area (POA) (a site of GnRH production) from unaffected animals of the *hpg* strain was transplanted into the anterior third ventricle of adult *hpg* mice. We report that in such implanted**

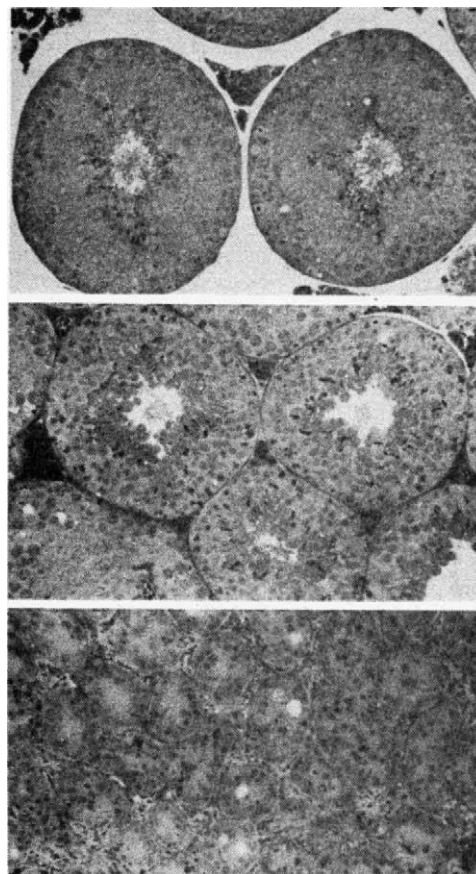
§ Present address: Neurology Department, Augustana Hospital, Chicago, Illinois 60611, USA.

animals, killed 2 months post-implantation, the POA grafts contained GnRH neurones, from which GnRH-positive fibres could be traced to capillaries of the median eminence. Hypothalamic GnRH and pituitary and plasma gonadotropin concentrations were increased compared with levels in untreated (*hpg*) animals. The testes were enlarged and had descended into the scrotum. Evidence of full spermatogenesis and interstitial cell development was present in testicular sections. No such effects were seen with transplants of cortical tissue.

Five experimental groups were established (see Fig. 1). All hosts were adult mice (*hpg* mutants, 5–6 months old; normal males, 5–8 months). POA grafts were obtained from 16–18-day-old fetuses (crown-to-rump length, 17–21 mm). (For method of grafting see Fig. 1 legend.) Growth rates of grafted animals were similar to those of unaffected mutants or normal animals. Two months post-implantation, blood was obtained for plasma gonadotropin determinations via retro-ocular puncture, then the animals were killed by decapitation. Testes and seminal vesicles were rapidly dissected and weighed and the testes were either placed in electron microscope fixative or flash-frozen. Brains were removed within 2 min: half of the brains of a given group were processed for immunocytochemical studies, the other half for GnRH immunoassay (see Fig. 1 legend).

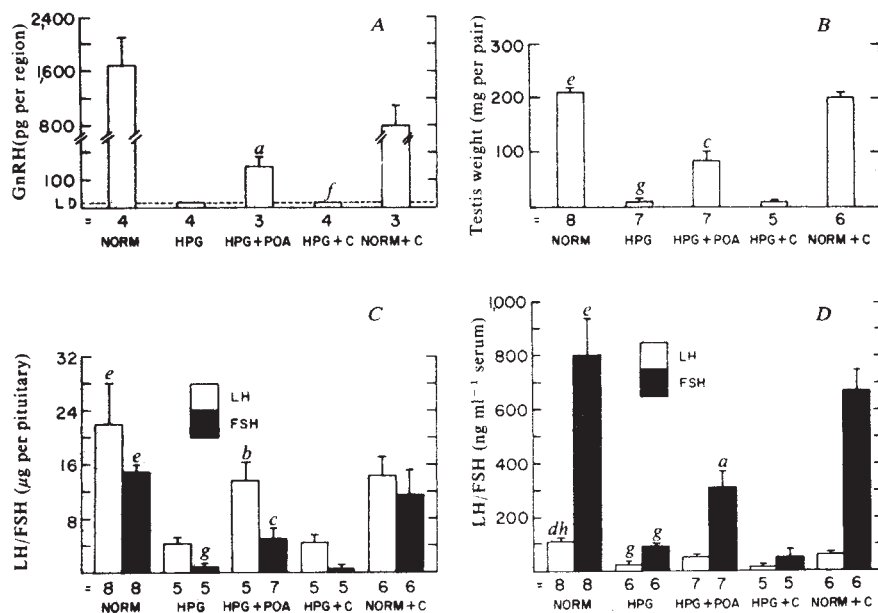
For histological studies, the brains from recipient animals were divided into three coronal blocks. The first coronal cut was made 2 mm in front of the optic chiasm, and the second behind the mammillary bodies. These blocks were fixed in Bouin's solution for 48 h, washed, embedded in paraffin, then cut serially at 6  $\mu$ m and placed on individual glass slides. Every 10th and 11th section was stained with cresyl violet and haematoxylin–eosin, respectively. Representative sections of each series were also immunoreacted for GnRH using a rabbit anti-GnRH serum (Nilaver–Silverman, no. III)<sup>3</sup> (see Fig. 3 legend).

Pituitary glands were bisected *in situ*; one half of each was frozen until the time of assay for gonadotropins, the other half placed overnight in Bouin's fixative for future immunocytochemical study. Murine luteinizing hormone (LH) and follicle-stimulating hormone (FSH) radioimmunoassays were performed according to the procedures recommended by



**Fig. 2** Testicular sections from normal animal (top); *hpg* animal with POA implant (middle); and *hpg* animal with cortex implant (bottom). Full spermatogenesis and evidence of interstitial cell development is present in POA recipients, while sections from animals with cortical implants are identical to those from untreated *hpg* animals.  $\times 154$ .

**Fig. 1** Effect of implants of preoptic area (POA) or cortex (C) on hypothalamic GnRH concentrations (A), testicular weight (B), and pituitary (C) and serum (D) gonadotropin concentrations. HPG, mutant hypogonadal mouse; NORM, normal littermate. Bars indicate mean  $\pm$  s.e.m. LD, limits of detection. Data (except for GnRH) were analysed using planned orthogonal comparisons; hypothalamic GnRH values were analysed with non-parametric methods due to unequal variances. We used Kruskal–Wallis one-way analysis of variance by ranks, followed by Ryan's procedure with  $\alpha = 0.05$  (ref. 17). GnRH values for *hpg* and *hpg* plus cortex groups were combined, as they were the same. POA–*hpg* compared with *hpg* and *hpg*–C:  $a = P < 0.05$ ;  $b = P < 0.02$ ;  $c = P < 0.01$ . NORM compared with *hpg*–POA:  $d = P < 0.01$ ;  $e = P < 0.001$ . NORM compared with *hpg*:  $f = P < 0.05$ ;  $g = P < 0.001$ . NORM compared with NORM–C:  $h = P < 0.05$ . The POA grafts were from 16–18-day-old fetuses. An anterior cut 1.0 mm wide was made at the midline of the ventral side of the fetal brain just caudal to the bifurcation of the anterior cerebral artery; a second cut was made  $\sim 0.7$  mm posterior to the first, just rostral to the hypothalamus. Lateral cuts were made 0.5 mm from the midline and the tissue block was finally cut 0.5 mm deep. The tissue was then placed in a drop of sterile saline in a Petri dish on a bed of ice and bisected with iris scissors. Two fetal preoptic tissue segments (comprising four pieces of tissue) were used per implant; cortical fragments of the same dimensions as the preoptic fragments were dissected from the frontal cortex of prepubertal mice. Graft recipients were anaesthetized with chloral hydrate and placed in a Kopf stereotaxic instrument. Grafts, in 2–4  $\mu$ l saline, were injected into the anterior third ventricle using a 22-gauge needle. Stereotaxic coordinates of the injection site were 5.5–6.0 mm down from the dura at the midline of the bregma. For assays for GnRH (ref. 18) after removal of the brain from recipients, the hypothalamus was dissected by coronal cuts 1 mm in front of the optic chiasm and behind the median eminence; laterally, to 1 mm lateral to the hypothalamic sulcus; and dorsally to the top of the third ventricle. It was placed in 0.5 ml of 0.1 M HCl. The remaining brain sections (excluding the brainstem from the midbrain down) were placed in 1 ml of 0.1 M HCl. Specimens were frozen until GnRH assay, at which time they were homogenized and neutralized.





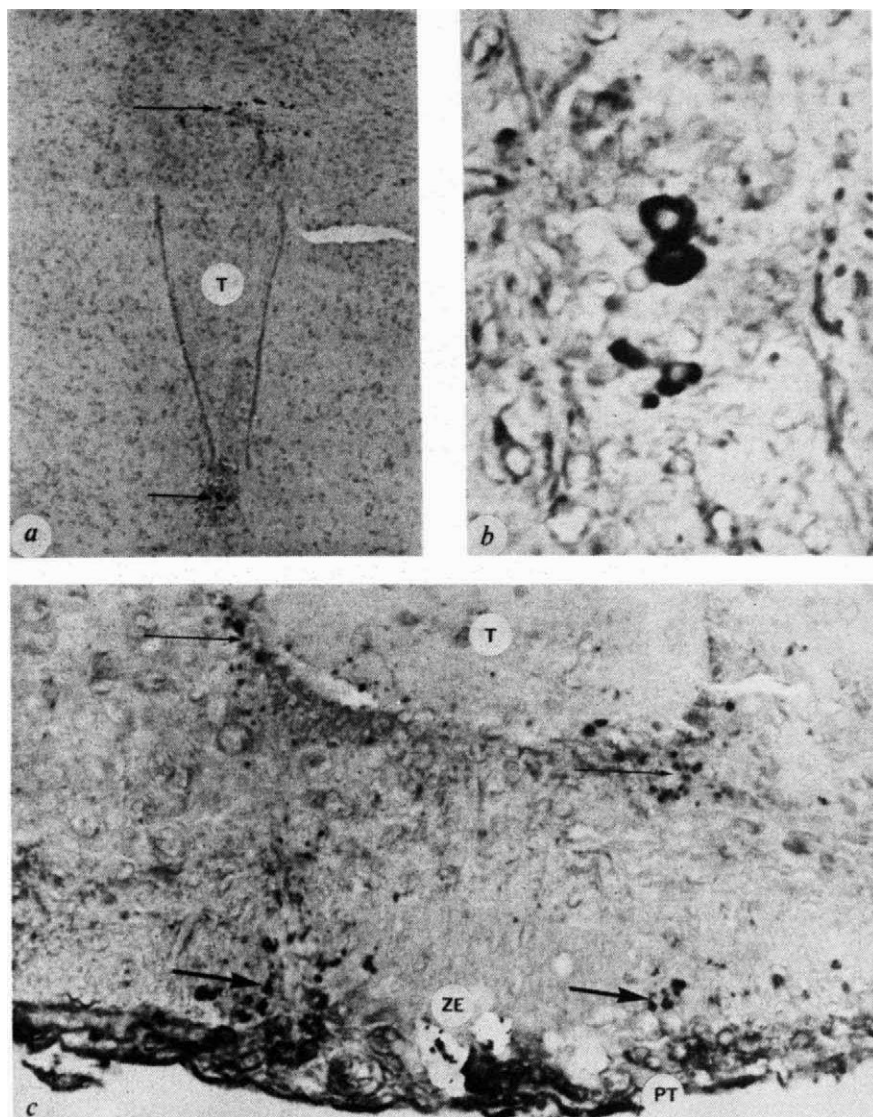
the NIAMDD Pituitary Hormone Distribution Program, with minor modifications, and used NIH-LH-S17 and NIH-FSH RP1 as standards.

Functional POA grafts were found in seven of eight implanted animals, as shown by significant increases in testicular size, pituitary and serum LH and FSH levels, and hypothalamic GnRH content, as compared with untreated animals or *hpg* animals having cortical transplants (Fig. 1). Plasma testosterone levels in the experimental animals were ( $\text{ng ml}^{-1}$ ): normal, 1.0; normal plus cortex, 1.7; *hpg* and *hpg* plus cortex, both undetectable; and *hpg* plus POA, 0.3. The increased plasma testosterone levels in the *hpg* plus POA group compared with the *hpg* group were compatible with the observed differences in seminal vesicle weight. These weights were (mg per pair): normal,  $258.3 \pm 25.6$ ; normal plus cortex,  $218.0 \pm 24.6$ ; *hpg*,  $17.2 \pm 1.6$ ; *hpg* plus cortex,  $14.0 \pm 2.9$ ; and *hpg* plus POA,  $66.1 \pm 16.2$ . In all groups combined, there were highly significant correlations between testis weight and hypothalamic GnRH, and between pituitary and serum gonadotropins. Pituitary and serum LH and FSH levels were highly correlated (all  $P < 0.01$ ). However, the correlation between pituitary LH and hypothalamic GnRH was not significant ( $r = 0.51$ ,  $n = 14$ ). Histological examination of the testes revealed full spermatogenesis in the POA recipients, while *hpg* animals with cortical implants were indistinguishable from those of untreated *hpg* animals (Fig. 2). Furthermore, the animals with the POA grafts showed evidence of interstitial cell development.

Studies of brain sections in the single animal without hormonal evidence of a functional transplant, revealed a vertical

cannula scar along the third ventricle which had destroyed part of the paraventricular and ventromedial nuclei on that side, with no evidence of transplanted tissue. There was no GnRH-immunoreactivity. In the remaining four POA-grafted animals that were examined anatomically, we found viable grafts containing GnRH neurones. In one animal, the graft grew as a cluster of cells within the ventrolateral hypothalamus on one side, and extended out under the median eminence to form an exophytic process under the pars tuberalis. GnRH-immunoreactive neuronal cell bodies were restricted to this cluster, while a few fibres were seen in the normal median eminence as well as in the cluster part of the transplant. In the other three animals, grafts were within the third ventricle; these extended ventrally into the median eminence. GnRH-reactive neuronal perikarya were present in the grafts; in all cases some fibres entered the median eminence in its lateral aspects, and others innervated capillaries in the zona externa (Fig. 3). No evidence of GnRH staining was found in untransplanted *hpg* animals. In normal adult mice, fibres in the median eminence were densely reactive, particularly in its lateral aspects, as previously reported<sup>4</sup>, and scattered immunoreactive cells were found in the rostral hypothalamus, including the medial preoptic area. Extrahypothalamic GnRH content ranged from 102 to 136 pg in control animals but was not detectable in any of the *hpg* animals (with or without grafts), also indicating that the POA grafts had innervated only the hypothalamic area.

Histological studies were also performed on three animals with cortical implants. In two of these, the injection sites were identified in the preoptic area on one side. In both cases, densely



**Fig. 3** Localization of GnRH in a *hpg* mouse hypothalamus containing a third ventricular transplant (T) of normal fetal mouse preoptic area. Deparaffinized brain sections were reacted with the primary anti-GnRH antiserum at 1:5,000 dilution for 12 h, followed by an avidin-biotin-peroxidase complex (ABC) method<sup>19</sup> using a Vectastain ABC kit (Vector Laboratories, California). After reaction with primary antiserum and buffer washes, biotinylated goat anti-rabbit IgG diluted to 1:200 was allowed to react with the tissue for 30 min, followed by avidin-HRP (horseradish peroxidase) for 1 h, then diaminobenzidine. *a*, Coronal section at the level of the mid-hypothalamus just rostral to the median eminence. The transplant fills the third ventricle and extends into the brain tissue just dorsal and ventral of it (arrows), where reactive neuronal perikarya were found. *b*, Higher magnification of three reactive neurones indicated by the lower arrow in *a*. *c*, Section more caudal to *a*, again showing the transplant within the third ventricle at the level of the mid-median eminence. Reactive fibres enter the dorsal median eminence at the ventrolateral corners of the graft (upper smaller arrows), and some can be traced to the zona externa (ZE) (lower, larger arrows). Some background staining can be seen in the pars tuberalis (PT) just external to the ZE. *a*,  $\times 112$ ; *b*,  $\times 764$ ; *c*,  $\times 360$ .



Nissl-reactive angular neurones resembling those normally found in cerebral cortex were widely distributed in the hypothalamus on the side of the injection, even extending into the thalamus and striatum. In the third animal, the injection had entered and widened the dorsal third ventricle. The transplanted cells, however, had not grown into the ventricle as in the case of the POA transplants. Again, transplanted cortical neurones widely and bilaterally infiltrated the areas described above. No GnRH-reactive perikarya were present in any of these animals.

Thus, partial restoration of hypothalamic GnRH levels in *hpg* animals by grafted GnRH cells apparently brings about the release of sufficient pituitary LH and FSH to cause greatly increased testicular weights and normal spermatogenesis and interstitial secretory activity. These results therefore demonstrate not only the presence of intact and functioning pituitary GnRH receptors, but also the receptivity of the undeveloped testicular tissue. Such findings are compatible with earlier reports of LH responsiveness (albeit diminished) to exogenous GnRH administration in *hpg* animals<sup>5</sup>, and of the normal function of *hpg* gonads when transplanted into normal animals<sup>6</sup>. It is possible that the increased hypothalamic GnRH in the POA-grafted mutant may have been secondary to restoration of a factor trophic to GnRH production that was lacking in the mutant, but this seems unlikely as the only GnRH cells seen were within the graft, not the host. The possibility that any neural tissue could stimulate GnRH production can also be discounted, as there were no hormonal or testicular changes in the *hpg* group with cortical implants.

There is evidence that by an extrapituitary action, GnRH potentiates sexual behaviour. This effect may occur via projections to the medial preoptic area, the arcuate nucleus, and mesencephalic central grey<sup>7</sup>. There are no data concerning restoration of normal sexual behaviour in the POA-grafted *hpg* animals. Although projections to the median eminence alone were seen in the POA grafts, it is possible that material secreted by the grafts into the CSF could reach the other areas mentioned.

The POA grafts did not restore hypothalamic GnRH to normal adult levels, although a high degree of restoration of pituitary and gonadal function was achieved. This raises interesting questions as to the relative contributions to total GnRH content of its synthesis as opposed to its release. It remains to be determined whether the reduced GnRH content in POA-grafted animals is a function of the amount of tissue implanted, or whether it is secondary to the lack of appropriate afferent connections to GnRH cells. Further immunocytochemical and electron microscopic studies are necessary to determine the nature of any such afferent connections to the grafted GnRH cells. It has been reported previously that host tissue can innervate fetal transplants<sup>8-11</sup>. Our findings of POA graft fibre outgrowths to blood vessels are similar to those reported by Stenevi *et al.*<sup>12</sup>, suggesting a trophic effect of such vessels on fibre outgrowth. Interestingly, the path of fibre outgrowth from such POA grafts was to the lateral areas of the median eminence, where GnRH fibres are found in the normal animal, whereas no such pathways were seen in outgrowths from cortical implants.

The present study extends previous observations with regard to correction of experimentally produced dopamine deficiency<sup>13,14</sup>, or of genetic vasopressin deficiency (in the Brattleboro rat)<sup>15,16</sup>. (In the latter study, only 5 of 40 transplants met criteria for establishing host responsiveness to transplantation of vasopressin neurones, and the study was carried out for only 20 days post-implantation.) The prolonged maintenance of normal hormone parameters and the high success rate achieved in the present study indicate that the *hpg*-grafted animal may provide a valuable model for studying the neural factors involved in the neuroendocrinology of gonadal function.

This work was supported in part by grants from the Lita Annenberg Hazen Fund, NIH grant HD-05077, Irma T. Hirschl

Trust, and the MRC (to H.M.C.). Drs T. Nett and N. G. Niswender provided anti-GnRH antiserum 42.

Received 18 February; accepted 11 May 1982.

1. Cattaneach, B. M., Iddon, C. A., Charlton, H. M., Chiappa, S. A. & Fink, G. *Nature* **269**, 338-340 (1977).
2. Björklund, A., Stenevi, U. & Svendgaard, N.-A. *Nature* **262**, 787-790 (1976).
3. Schwanzel-Fukuda, M. & Silverman, A. J. *J. comp. Neurol.* **191**, 213-225 (1980).
4. Gross, D. S. & Baker, B. L. *Am. J. Anat.* **148**, 195-216 (1977).
5. Iddon, C. A., Charlton, H. M. & Fink, G. *J. Endocr.* **85**, 105-110 (1980).
6. Bamber, S., Iddon, C. A., Charlton, H. M. & Ward, B. J. *J. Reprod. Fert.* **58**, 249-252 (1980).
7. Moss, R. L., Riskind, P. & Dudley, C. in *Central Nervous System Effects of Hypothalamic Hormones and Other Peptides* (eds Collu, R., Barbeau, A., Ducharme, J. R. & Rochefort, J.-G.), 345 (Raven, New York, 1979).
8. Lund, R. D. & Hauschka, S. D. *Science* **193**, 582-584 (1976).
9. Hallas, B. H., Oblinger, M. M. & Das, G. D. *Brain Res.* **196**, 242-246 (1980).
10. Kromer, L. F., Björklund, A. & Stenevi, U. *Brain Res.* **210**, 153-171 (1981).
11. Azmitia, E. C., Perlow, M. J., Brennan, M. J. & Lauder, J. M. *Brain Res. Bull.* **7**, 703-710 (1981).
12. Stenevi, U. *et al. Cell Tissue Res.* **205**, 217-228 (1980).
13. Perlow, M. J. *et al. Science* **204**, 643-647 (1979).
14. Björklund, A., Donnett, S. B., Stenevi, U., Lewis, M. F. & Iversen, S. D. *Brain Res.* **199**, 307-333 (1980).
15. Gash, D. & Scott, D. E. *Cell Tissue Res.* **211**, 191-206 (1980).
16. Gash, D., Sladek, J. R. Jr & Sladek, C. *Science* **210**, 1367-1369 (1981).
17. Kirk, R. F. *Experimental Design: Procedures for the Behavioral Sciences*, 69-98; 493-497 (Wadsworth, California, 1968).
18. Araki, S., Ferin, M., Zimmerman, E. A. & vande Wiele, R. I. *Endocrinology* **96**, 644-650 (1975).
19. Hsu, S.-M., Raine, L. & Fanger, H. J. *Histochem. Cytochem.* **29**, 577-580 (1981).

## Shiverer peripheral myelin contains P<sub>2</sub>

Janet Winter

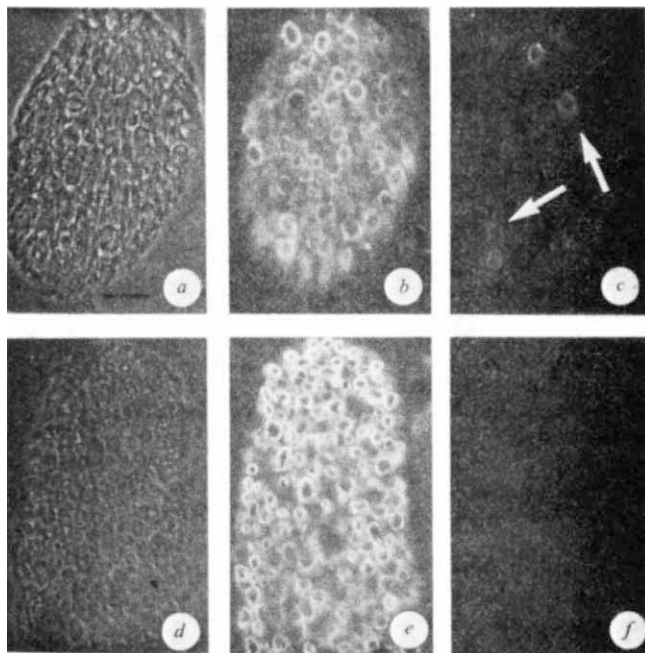
MRC Neuroimmunology Project, Department of Zoology, University College London, Gower Street, London WC1E 6BT, UK

Myelin-deficient mutant mice, such as shiverer, can provide information about the normal mechanisms involved in myelination. The shiverer mouse carries a recessive, autosomal mutation<sup>1</sup> resulting in an extreme deficiency in central myelin, and the small amount of myelin present is poorly compacted<sup>2</sup>; the peripheral myelin, however, appears essentially normal<sup>3,4</sup>. As the amount of myelin basic protein (P<sub>1</sub>) in both central and peripheral nervous system myelin is extremely low in shiverer<sup>3,5</sup>, it is possible that P<sub>1</sub> is essential for the normal formation and compaction of central myelin, but not of peripheral myelin. Some other protein would then be responsible for the formation of compact peripheral myelin in shiverer. Peripheral myelin contains another basic protein, designated P<sub>2</sub>, which could be a possible candidate for this role. Kirschner and Ganser, however, using SDS-polyacrylamide gel electrophoresis, reported that P<sub>2</sub>, as well as P<sub>1</sub>, is absent from shiverer sciatic nerve<sup>3</sup>. This is an important observation if correct, because it not only excludes the possibility that P<sub>2</sub> is required for compaction but also makes it less likely that the deficiency in P<sub>1</sub> is the primary defect in shiverer. As P<sub>2</sub> in rat and mouse has frequently been confused with another small basic protein (related to P<sub>1</sub>) in SDS-polyacrylamide gels<sup>6</sup>, it seemed worthwhile to reassess this aspect of the Kirschner and Ganser observations. Immunohistochemistry and immunoblotting have been used here to show unambiguously that P<sub>2</sub> is present in shiverer peripheral myelin.

By immunofluorescence the uneven morphological distribution of P<sub>2</sub> in peripheral myelin sheaths<sup>7,8</sup> was the same in shiverer and normal mice (Fig. 1). Immunoblotting revealed in the shiverer a band of the same mobility as bovine P<sub>2</sub> and normal mouse P<sub>2</sub>, in about the same amounts as in normal mouse myelin (Fig. 2). The major myelin peripheral glycoprotein P<sub>0</sub> was also distributed in normal fashion in the shiverer, whereas P<sub>1</sub> was absent.

The presence of P<sub>2</sub> in shiverer peripheral myelin is perhaps not surprising. The shiverer is reportedly a one-gene mutation and therefore only one gene product should be directly affected.



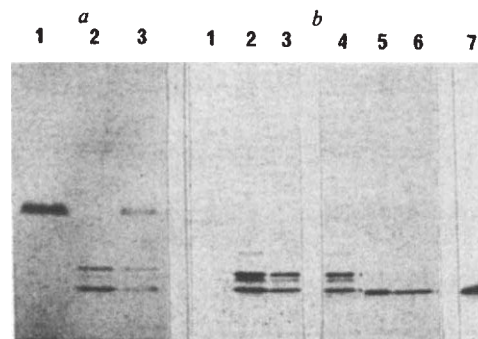


**Fig. 1** Antisera used were rabbit anti-rat P<sub>0</sub>, (R anti-P<sub>0</sub>), rabbit anti-bovine basic protein (R anti-bp) and rabbit anti-bovine P<sub>2</sub> (R anti-P<sub>2</sub>). Tetramethyl rhodamine and fluorescein isothiocyanate conjugated to goat anti-rabbit immunoglobulins (G anti-R rhodamine, G anti-R fluorescein) were obtained from Nordic Laboratories. Sections (8 μm) of formaldehyde-fixed tissue were cut on a Brighton cryostat and mounted on dry, gelatin-coated microscope slides. All sections were ethanol-extracted for 10 min and washed in phosphate-buffered saline (PBS) before treatment with either R anti-P<sub>2</sub> or R anti-bp diluted 1:50 in PBS (30 min), then washed and incubated with G anti-R rhodamine (1:150). After further washing, R anti-P<sub>0</sub> diluted 1:20 was applied, washed and followed by G anti-R fluorescein (1:50)<sup>7</sup>. (As controls, some sections were incubated with normal rabbit serum instead of specific antibodies.) After a final wash, sections were mounted in 50% glycerol and viewed using a Zeiss Universal microscope equipped with both phase and fluorescence optics. *a*, Phase contrast of 8-μm section of dorsal roots from the shiverer mouse. *b*, P<sub>0</sub> staining visualized by R anti-P<sub>0</sub> followed by G anti-R fluorescein. Note the even staining of myelin sheaths. *c*, P<sub>2</sub> staining visualized by R anti-P<sub>2</sub> followed by G anti-R rhodamine. Note that some sheaths are stained much more strongly than others (marked by arrows). *d*, Phase contrast of 8-μm section of dorsal roots from shiverer mouse. *e*, P<sub>0</sub> staining visualized by R anti-P<sub>0</sub> followed by G anti-R fluorescein. *f*, bp staining, R anti-bp followed by G anti-R rhodamine. Note the lack of staining. All control sections were unstained. Scale bar, 30 μm.

As P<sub>2</sub> is unrelated to basic protein<sup>9</sup>, a mutation in the basic protein gene should not affect P<sub>2</sub>. Although Kirschner and Ganser reported an absence of P<sub>2</sub> on SDS polyacrylamide gel electrophoresis of shiverer peripheral nerve, they were probably mistaking the small basic protein for P<sub>2</sub> (ref. 6). In fact, P<sub>2</sub> runs slightly ahead of this small basic protein and is present in such small quantities that it is usually not visualized with Coomassie blue, although Mikoshiba *et al.* state that they can see P<sub>2</sub> on stained gels of shiverer peripheral myelin<sup>10</sup>. It is easily detected with the more specific and sensitive immunautoradiographic methods used in the present study.

It seems unlikely that P<sub>2</sub> functions in peripheral myelin in the same way as basic protein does in central myelin for two reasons. It is present in extremely small quantities in normal peripheral myelin, whereas basic protein normally constitutes up to 30% of normal central myelin. Furthermore, P<sub>2</sub> seems to be virtually absent in some peripheral myelin sheaths, which in both shiverer and normal mice show normal compact myelin formation. Thus, the function of P<sub>2</sub> remains unknown.

I thank Rhona Mirsky and Martin Raff for helpful advice and discussion, Robert I. Craggs for supplying the shiverer



**Fig. 2** Purified bovine P<sub>2</sub> was used as a standard. Normal mouse central myelin was prepared from brain<sup>11</sup>, and normal and shiverer peripheral myelin was purified from sciatic nerve by the method of Keenan and Jones<sup>12</sup> except that the nerve was initially sonicated in sucrose solution. Purified myelin was lyophilized, extracted with ethanol/ether (3/2), then solubilized by boiling in 2% SDS. Total protein content was estimated by the method of Lowry<sup>13</sup>. Samples were run on 10-15% gradient Laemmli gels<sup>14</sup> and then electrophoretically transferred to 0.12-μm pore size nitrocellulose paper (Schleicher and Schüll)<sup>15</sup>. The paper was then incubated for 1 h with 5% haemoglobin (Sigma bovine type 11), followed by either R anti-P<sub>2</sub> or R anti-bp diluted 1:1,000 in 5% haemoglobin in PBS, the paper was incubated with iodinated protein 'A' (Sigma) (~10<sup>7</sup> c.p.m. in 20 ml 5% haemoglobin PBS) for 4 h at room temperature. The nitrocellulose papers were then washed thoroughly in PBS, dried and autoradiographed overnight on Kodak X-ray film. *a*: Coomassie-blue stained Laemmli gel (10-15%). 1, 25 μg shiverer peripheral myelin; 2, 15 μg normal central myelin; 3, 25 μg normal peripheral myelin. *b*: 1, 25 μg shiverer peripheral myelin; 2, 15 μg normal central myelin; 3, 25 μg normal peripheral myelin; 4, 25 μg normal peripheral myelin; 5, 25 μg normal peripheral myelin; 6, 25 μg shiverer peripheral myelin; 7, 0.1 μg purified bovine P<sub>2</sub>. Tracks 1-3 were nitrocellulose transfers from a polyacrylamide gel and were treated with R anti-bp followed by protein A. Tracks 4-6 were transfers from the same gel. This time, track 4 was cut carefully, the left-hand portion stained with R anti-bp, but the right hand portion and tracks 5 and 6 were incubated with R anti-P<sub>2</sub>. After application of <sup>125</sup>I-labelled protein A to the papers, and washing and drying, the tracks were carefully lined up before autoradiography. Tracks 1-6 were exposed on Kodak X-ray film for 12 h whereas track 7 was exposed for 4 h only. The normal myelin showed the four basic protein bands<sup>16</sup> (in both central and peripheral myelin), whereas no bands of this mobility were visualized in the shiverer peripheral myelin, in agreement with earlier reports<sup>3</sup>.

mice, Jeremy Brockes for gifts of rabbit anti-rat P<sub>0</sub> antiserum, Richard Hughes and Michael Kadlubowski for gifts of rabbit anti-bovine P<sub>1</sub> and P<sub>2</sub> antisera and purified bovine P<sub>2</sub>, and James Cohen for help with photography. I was supported by a SERC research studentship.

Received 5 March; accepted 10 May 1982.

1. Bird, T., Farrel, D. F. & Sumi, S. H. *Trans. Am. Soc. Neurochem.* **8**, 153 (1977).
2. Privat, A., Jaque, C., Bourre, J. M., Dupouey, P. & Baumann, N. *Neurosci. Lett.* **12**, 107-112 (1979).
3. Kirschner, D. A. & Ganser, A. L. *Nature* **283**, 207-210 (1980).
4. Rosenbluth, J. J. *comp. Neurol.* **193**, 729-739 (1980).
5. Dupouey, P., Jaque, C., Bourre, J. M., Cesselin, F., Privat, A. & Baumann, N. *Neurosci. Lett.* **12**, 113-118 (1979).
6. Greenfield, S., Brostoff, S. W. & Hogan, E. L. *J. Neurochem.* **34**, 453-455 (1980).
7. Winter, J., Mirsky, R. & Kadlubowski, M. *J. Neurocytol.* **11**, 351-362 (1982).
8. Trapp, B., McIntyre, L. J., Quarles, R. H., Sternberger, N. H. & Webster, H. *DeF. Proc. nam. Acad. Sci. U.S.A.* **76**, 3552-3556 (1979).
9. Brostoff, S. W., Kakhams, Y. D., Carlo, D. J., Reuter, W. & Eylar, E. H. *Brain Res.* **86**, 449-458 (1975).
10. Mikoshiba, K., Kohsaka, S., Takamatsu, K. & Tsukada, Y. *Brain Res.* **204**, 455-460 (1981).
11. Norton, W. T. & Poduslo, S. E. *J. Neurochem.* **21**, 749-758 (1973).
12. Keenan, E. W. & Jones, M. J. *Neurochem.* **30**, 231-235 (1980).
13. Lowry, O. H., Rosebrough, M. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
14. Laemmli, U. K. & Favre, M. *J. Molec. Biol.* **80**, 575-600 (1973).
15. Vaessen, R. T. M. J., Kreike, J. & Groot, G. S. P. *FEBS Lett.* **124**, 193-196 (1981).
16. Barhause, E., Braun, P. E. & Carson, J. H. *Proc. nam. Acad. Sci. U.S.A.* **74**, 3360-3364 (1977).

# Modulation of *in vitro* BFU-E growth by normal Ia-positive T cells is restricted by HLA-DR

Beverly Torok-Storb & John A. Hansen

Fred Hutchinson Cancer Research Center, and the Puget Sound Blood Center, Seattle, Washington 98104, and the Department of Medicine, Division of Oncology, University of Washington, School of Medicine, Seattle, Washington 9815, USA

The *in vitro* growth of immature erythroid precursors, assayed as burst forming units (BFU-E), can be enhanced by co-culturing with normal T cells<sup>1-3</sup>. The growth enhancing effect of normal T cells, however, was shown to be the outcome of at least two functionally distinct subpopulations, one that enhances burst formation and one that limits it<sup>3,4</sup>. It appears that the T cell population limiting growth enhancement of the erythroid precursors expresses 'Ia-like' antigen (Ia<sup>+</sup>) because treatment of T cells with anti-Ia antibody and complement significantly increased the ability of the remaining T cells to stimulate autologous BFU-E growth<sup>5</sup>. As BFU-E also seem to express Ia, the inhibiting effect of certain T cells on BFU-E growth may involve the Ia molecule. We show here that the inhibitory effect of Ia<sup>+</sup> T cells on BFU-E growth is genetically restricted to cell combinations which are phenotypically identical for at least one HLA-DR (Ia-like) antigen.

The presence of Ia antigen on the surface of erythroid precursors was determined by using a murine monoclonal antibody, called 7.2, which is specific for a common determinant on the human 'Ia-like' p29, 34 bimolecular complex<sup>5</sup>. In the experiments summarized in Table 1, two-stage complement (C') mediated cytotoxicity was used to deplete Ia<sup>+</sup> cells from normal bone marrow and peripheral blood. The data presented here indicate that growth of immature precursors derived from blood or bone marrow was significantly reduced by antibody and C' whereas growth of the more mature precursors, day 5 CFU-E, was not. These data are in agreement with the findings of Robinson *et al.* who used fluorescence activated cell sorting to collect Ia<sup>+</sup> bone marrow cells which contained the BFU-E, but not the CFU-E population<sup>6</sup>. In studies presented elsewhere we have used positive selection experiments to demonstrate that BFU-E can be found in Ia<sup>+</sup> peripheral blood mononuclear cells. This was necessary to rule out the possibility that the reduced growth of colonies observed following antibody and C' treatment was due to removal of critical Ia<sup>+</sup> auxiliary cells.

Studies of marrow derived erythroid precursors, summarized in Table 1, indicate that Ia antigens are expressed on immature BFU-E but are gradually lost as these cells mature. It is obvious from these and other data that marrow samples contain a very

**Table 2** Effect of T cells on BFU-E from monocyte depleted non-T cells

| T cells added           | % Control BFU-E growth from non-TPBM (mean ± s.e.) |                   |
|-------------------------|----------------------------------------------------|-------------------|
|                         | Monocytes present                                  | Monocytes removed |
| Control T cells         | 180 ± 16                                           | 180 ± 37          |
| Ia <sup>-</sup> T cells | 280 ± 4                                            | 150 ± 42          |

Results are expressed as % of control BFU-E growth (no added T cells) and represent the mean ± s.e. of 10 separate experiments. T cells were collected from Ficoll-Hypaque separated PBM by rosetting with 2-aminoethyl-isothiocyanate bromide (AET) treated sheep red blood cells. Monocytes were depleted by either plastic adherence or treatment with murine monoclonal anti-monocyte antibodies and C'. Antibody plus C' treatments were designed to assure complete removal of monocytes and are described in Table 1. Antibodies used were developed in this institution and are designated 5F1 (ref. 9) and 20.2 (20.2 is a murine monoclonal antibody which reacts with all peripheral blood and bone marrow monocytes, unpublished data). Both BFU-E sources (non-T PBM with or without monocytes) were cultured alone for control BFU-E growth and with T cells treated with C' only (control T cells) or with 7.2 and C' (Ia<sup>-</sup> T cells). 1 × 10<sup>5</sup> non-T cells were cultured with 4 × 10<sup>5</sup> T cells per ml in plasma clots. Control BFU-E growth obtained from cultures without T cells ranged from 15 ± 2 to 61 ± 9 BFU-E per 10<sup>5</sup> non-T cells. Comparable numbers of BFU-E were obtained from 10<sup>5</sup> monocyte-depleted non-T cells.

heterogenous population of precursors representing several maturational stages, and that each of these may be subject to different forms of regulation. In comparison, circulating BFU-E represent a relatively homogeneous class of immature precursors of which greater than 90% are Ia<sup>+</sup>. For this reason and also because T cells have a more pronounced effect on growth of blood derived BFU-E compared to marrow derived BFU-E<sup>7</sup>, we limited the current study to T cell effects on precursors grown from non-T peripheral blood mononuclear cells (PBM).

The non-T cell population from which BFU-E are grown contains not only Ia<sup>+</sup> BFU-E, but also Ia<sup>+</sup> monocytes. As monocytes are involved in many T cell functions<sup>8</sup> we investigated their possible role in T cell modulation of BFU-E growth. Table 2 summarizes data from ten experiments in which control T cells were cultured with autologous non-T cells in the presence or absence of monocytes. The data indicate that the presence or absence of monocytes has no effect on the growth enhancing ability of autologous control T cells. In contrast, the increased stimulating activity of Ia<sup>-</sup> T cells is found only when monocytes are present. This indicates that the increased BFU-E enhancing activity of Ia<sup>-</sup> T cells is monocyte dependent.

Table 3 summarizes data from 24 experiments in which T cell depleted peripheral blood mononuclear cells (PBM) were cultured for BFU-E growth with and without added T cells. In all cases, control T cells treated with medium and C' increased

**Table 1** Effect of anti-Ia and C' treatment on erythroid precursors

| Cell source      | No. of experiments | Precursor assayed | % Control growth (mean ± s.e.) |
|------------------|--------------------|-------------------|--------------------------------|
| Bone marrow      | 8                  | Day 5 CFU-E       | 81 ± 11.6                      |
| Bone marrow      | 5                  | Day 10 BFU-E      | 44 ± 18.4                      |
| Bone marrow      | 8                  | Day 15 BFU-E      | 18 ± 3.9                       |
| Peripheral blood | 10                 | Day 14 BFU-E      | 8 ± 2.1                        |

Bone marrow buffy coat cells and Ficoll-Hypaque separated peripheral blood mononuclear cells were suspended at a concentration of 1–4 × 10<sup>6</sup> in 0.1 ml tissue culture medium (RPMI 1640) containing 20% fetal bovine serum. An equal volume of 7.2 ascites fluid diluted 1/1,000 was added and the cells incubated for 30 min at room temperature. Pooled normal rabbit serum (0.2 ml) was added as a source of C'. Incubation was continued at room temperature for 60 min, cells were washed three times, resuspended to starting volume and then assayed for colony growth in plasma clots. Control values were obtained from cells treated with C' only.

**Table 3** Effect of control and Ia<sup>-</sup> T cells on BFU-E

| Cell combinations                       | No. of experiments | % Control BFU-E growth (mean ± s.e.) |                         |            |
|-----------------------------------------|--------------------|--------------------------------------|-------------------------|------------|
|                                         |                    | Control T cells                      | Ia <sup>-</sup> T cells | % Increase |
| Autologous                              | 24                 | 200 ± 22                             | 342 ± 53                | 71*        |
| Parent-child one DR haplotype identical | 6                  | 260 ± 47                             | 498 ± 130               | 91*        |
| Unrelated, DR identical                 | 6                  | 183 ± 50                             | 357 ± 100               | 96*        |
| Unrelated, DR non-identical             | 12                 | 250 ± 64                             | 225 ± 66                | 0          |

T cells were separated from peripheral blood mononuclear cells by rosetting with AET treated sheep red blood cells. Non-rosetting cells, designated non-T, were cultured alone for control BFU-E growth and with T cells treated either with C' only (control T cells) or with 7.2 and C' (Ia<sup>-</sup> T cells). 2 × 10<sup>5</sup> non-T cells were cultured with 4 × 10<sup>5</sup> T cells/ml in plasma clots. Control BFU-E growth in the 24 experiments ranged from 11 ± 2 to 58 ± 5/10<sup>5</sup> non-T cells.

\* P ≤ 0.001.



**Table 4** Mixed leukocyte culture responses of cell combinations used in co-culture experiments

| Cell combinations                        | No. of experiments | MLC reactivity |                       |
|------------------------------------------|--------------------|----------------|-----------------------|
|                                          |                    | c.p.m.         | Relative response (%) |
| Autologous                               | 40                 | 1,004 ± 157    | 0                     |
| Parent-child, one DR haplotype identical | 6                  | 40,365 ± 8,478 | 68 ± 11               |
| Unrelated, DR identical                  | 20                 | 9,225 ± 1,909  | 23 ± 4.9              |
| Unrelated, DR nonidentical               | 4                  | 49,145 ± 4,442 | 80 ± 3.7              |

Phenotype compatibility for HLA-DR was assessed only for individuals expressing DR 1-4 or 7, and identity versus nonidentity was defined only for noncrossreacting antigens. MLC reactivity is expressed as counts per minute (c.p.m.) in triplicate cultures (mean ± s.e.) and % relative response (the response to a pool of stimulator cells from 4 unrelated donors is defined as 100%).

BFU-E growth as compared to cultures without T cells. This was true for all combinations of cells whether autologous or from allogeneic donors regardless of DR compatibility. In contrast, the increased stimulating ability of Ia<sup>+</sup> T cells was only observed with cell combinations sharing at least one DR antigen. In twelve experiments using cells from DR phenotypically identical or haplo-identical donors, T cell stimulating ability was significantly increased after depletion of Ia<sup>+</sup> cells. In contrast, in all twelve experiments using cell combinations incompatible for both DR haplotypes the opposite effect was seen. Following depletion of Ia<sup>+</sup> T cells, stimulating ability was slightly decreased. In all experiments, T cells were also cultured with autologous non-T cells. In each case Ia<sup>+</sup> T cells showed a significant increase in their ability to stimulate autologous BFU-E.

These data suggest that the ability of Ia<sup>+</sup> T cells to limit the growth of Ia<sup>+</sup> BFU-E was DR restricted. The effect was not related to the level of reactivity in MLC (Table 4). Autologous cells are nonreactive in MLC, while one haplotype identical parent-child combinations are vigorously reactive. Both cell combinations however show increased stimulating ability after Ia<sup>+</sup> T cell depletion.

Demonstration of DR restriction in this system suggests that the human 'Ia-like' molecule may function as a receptor in an interaction between the growth limiting Ia<sup>+</sup> T cell and the BFU-E to limit the responsiveness of BFU-E to monocyte dependent, Ia<sup>+</sup> T cell stimulating activity. This inhibitory activity of the Ia<sup>+</sup> T cell on BFU-E may be direct, or the Ia<sup>+</sup> T cell may interact with a monocyte to prevent Ia<sup>+</sup> T cell stimulating activity. The first hypothesis would require DR restriction between T cells and BFU-E, the second would require restriction between T cells and monocytes. So far, attempts to distinguish between these two possibilities have not been successful because manipulation of monocytes results in their nonspecific activation. Nevertheless, our data indicate DR restriction of a T cell regulatory mechanism that limits in vitro erythropoiesis, an observation that implicates the Ia molecule in a new biological role.

We thank Agnes Chang for technical assistance and Pauline Marsden for preparing the manuscript. Antibodies 7.2 and 20.2 were provided by Drs Paul J. Martin and Malek Kamoun and antibody 5F1 by Dr Irwin D. Bernstein. This investigation was supported by NCI grants CA-18221, CA-18029, CA-29548, CA-15704 and the Heart, Lung and Blood Institute grants HL-17265 and HL-24439. BT-S is a Special Fellow of the Leukemia Society of America.

Received 13 April; accepted 27 May 1982.

1. Nathan, D. G. *et al.* *J. exp. Med.* **147**, 324-339 (1978).
2. Mangan, K. F. & Desforages, J. F. *Exptl Hematol.* **8**, 717-727 (1980).
3. Torok-Storb, B., Martin, P. J. & Hansen, J. A. *Blood* **58**, 171-174 (1981).
4. Torok-Storb, B., Martin, P. J. & Hansen, J. A. in *Hemoglobin in Development and Differentiation* (eds Stamatoyannopoulos, G. & Nienhuis, A. W.) 137-143 (Liss, New York, 1981).

5. Hansen, J. A., Martin, P. J. & Nowinski, R. C. *Immunogenetics* **10**, 247-260 (1980).
6. Robinson, J., Sieff, C., Delia, D., Edwards, P. A. W. & Greaves, M. *Nature* **289**, 68-71 (1981).
7. Lipton, J. M. *et al.* *J. exp. Med.* **152**, 350-360 (1980).
8. Winchester, R. J. & Kunkel, H. G. *Adv. Immun.* **28**, 221-292 (1979).
9. Bernstein, I. D., Andrews, R. G., Cohen, S. F. & McMaster, B. E. *J. Immun.* **128**, 876-881 (1982).

## Correlation between immunoglobulin light chain expression and variant translocation in Burkitt's lymphoma

G. M. Lenoir\*, J. L. Preud'homme†, A. Bernheim‡ & R. Berger‡

\*International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 2, France

†Laboratory of Immunochimistry and Immunopathology, INSERM U 108, Hôpital St Louis, 75475 Paris Cedex 10, France and Laboratory of Immunology and Immunopathology, CHUR la Milétrie, 86021 Poitiers Cedex, France

‡Laboratory of Cytogenetics, Institut de Recherches sur les Leucémies et Maladies du Sang, UER d'Hématologie, Hôpital St Louis, 2 place du Dr Fournier, 75475 Paris Cedex 10, France

Burkitt's-type lymphomas-leukaemias (BL) are monoclonal proliferations of malignant B lymphocytes<sup>1</sup>. Irrespective of whether they carry the Epstein-Barr virus (EBV) genome, these tumour cells have been shown consistently to have one of the specific reciprocal chromosome translocations, t(8; 14), t(2; 8) or t(8; 22), involving the long arm of chromosome 8 (on 8q24) and chromosome 14, 2 or 22 (on 14q32, 2p12 and 22q11, respectively)<sup>2-6</sup>. The latter chromosomes have been shown recently to carry genes for immunoglobulin (Ig) heavy chains<sup>7,8</sup>, and  $\kappa$ <sup>9</sup> and  $\lambda$ <sup>10,11</sup> light chains, respectively. Furthermore, the localization of  $\kappa$  light chains within 2pcen-2p13 encompasses the breakpoint observed in Burkitt's translocation (2p12). It was therefore considered of interest to determine whether the expression of immunoglobulin chains in BL cells is related to the type of chromosomal anomalies observed. We report here that there is a direct relationship between expression of immunoglobulin light chains and specific type of translocation: BL cells with t(8; 22) express  $\lambda$  chains, whereas those with t(2; 8) express  $\kappa$  chains.

We studied immunoglobulin expression in BL cells in 19 continuous BL cell lines, and in 10 fresh specimens. For the immunological and cytogenetic investigations, all specimens were coded and studied 'blind'. The study was performed at least twice for most of the samples. In two cases, the fresh material and continuous cell lines gave identical results. Table 1 summarizes the main characteristics of the lymphoid lines used: 18 were derived from Burkitt malignant lymphomas and one from a Burkitt-type leukaemia. Most (12/19) of the lines were derived from cases originating in areas considered to be non-endemic for this tumour, and 7 did not contain the EBV genome; 13 of the lines were established at the International Agency for Research on Cancer. The 10 cases from which fresh material was obtained were 9 Burkitt's-type leukaemias and 1 Burkitt's lymphoma. All patients were investigated at the Hôpital St Louis, Paris. Chromosome analysis was performed using GTG and RTG banding techniques<sup>12</sup>, and the karyotypes were established following international nomenclature<sup>12</sup>. Surface and intracytoplasmic immunoglobulin (SIg and cIg) were detected by immunofluorescent staining using conjugated F(ab')<sub>2</sub> fragments of rabbit IgG anti-human  $\mu$ ,  $\gamma$ ,  $\alpha$ ,  $\delta$ ,  $\lambda$  and  $\kappa$  chains, as described previously<sup>13</sup>.

Of the 19 cell lines used, 9 were shown to contain a t(8; 14) translocation, 5 a t(8; 22) and 5 a t(2; 8) (see Table 1). Nine of the fresh specimens had t(8; 14) and one had t(8; 22). With regard to immunoglobulin expression, the phenotype of the BL cells varied from a pre-B (four lines expressed predominantly

Table 1 Immunoglobulin expression in Burkitt's lymphoma-leukaemia cells

| Specimen              | Sex/Age | Origin | EBV     | Trans-<br>location | SIg*        |             |          | cIg*        |           |          | Fc<br>receptor* |     |    |      |
|-----------------------|---------|--------|---------|--------------------|-------------|-------------|----------|-------------|-----------|----------|-----------------|-----|----|------|
|                       |         |        |         |                    | Heavy chain | $\lambda$   | $\kappa$ | Heavy chain | $\lambda$ | $\kappa$ |                 |     |    |      |
| Continuous cell lines |         |        |         |                    |             |             |          |             |           |          |                 |     |    |      |
| IARC/BL 3             | M       | 09     | France  | —                  | t(8;14)     | $\mu$       | 6        | 3           | 0         | $\mu$    | 47              | 20  | 0  | 0    |
| IARC/BL 8             | M       | 03     | Algeria | +                  | t(8;14)     | $\mu$       | 100      | 100         | 0         | $\mu$    | 15              | 15  | 0  | 100  |
| IARC/BL 14            | F       | 36     | Turkey  | —                  | t(8;14)     | $\mu$       | 14       | 0           | 34        | $\mu$    | <0.5            | 0   | 2  | 0    |
| IARC/BL 28            | M       | 19     | France  | —                  | t(8;14)     | $\mu$       | 15       | 0           | 0         | $\mu$    | 80              | 0   | 0  | 0    |
| IARC/BL 29            | F       | 04     | Reunion | +                  | t(8;14)     | $\mu$       | 100      | 0           | 100       | $\mu$    | 0               | 0   | 0  | 0    |
| IARC/BL 31            | M       | 14     | France  | —                  | t(8;14)     | $\mu$       | 12       | 0           | 58        | $\mu$    | 19              | 0   | 57 | 0    |
| CHEV                  | M       | 14     | France  | —                  | t(8;14)     | $\mu$       | 100      | 0           | 100       | ND       | ND              | ND  | ND | ND   |
| DAUDI                 | M       | 16     | Kenya   | +                  | t(8;14)     | $\mu$       | 100      | 0           | 100       | $\mu$    | 0               | 0   | 0  | 70   |
| RAJI                  | M       | 12     | Nigeria | +                  | t(8;14)     | $\mu$       | 2        | 2           | 0         | $\mu$    | 88              | 88  | 0  | 0    |
| IARC/BL 2             | M       | 07     | France  | —                  | t(8;22)     | $\mu$       | 3        | 0           | 0         | $\mu$    | 92              | 72  | 0  | 0    |
| IARC/BL 33            | M       | 35     | France  | —                  | t(8;22)     | $\mu$       | 91       | 91          | 0         | $\mu$    | 100             | 100 | 0  | —0.5 |
| IARC/LY 47            | M       | ?      | Uganda  | +                  | t(8;22)     | $\mu$       | 2        | 0           | 0         | $\mu$    | 100             | 0   | 0  | 0    |
| IARC/LY 67            | M       | 08     | Kenya   | +                  | t(8;22)     | $\mu$       | 77       | 77          | 0         | $\mu$    | 85              | 70  | 0  | 77   |
| MAKU                  | M       | ?      | Kenya   | +                  | t(8;22)     | $\mu$       | 32       | 32          | 0         | $\mu$    | 53              | 53  | 0  | 0    |
| IARC/BL 21            | M       | 09     | Algeria | +                  | t(2;8)†     | $\mu$       | 100      | 0           | 100       | $\mu$    | 0               | 0   | 0  | 100  |
| IARC/LY 66            | M       | 13     | Kenya   | +                  | t(2;8)      | $\mu$       | 4        | 0           | 0         | $\mu$    | 96              | 0   | 0  | 0    |
| IARC/LY 91            | F       | 07     | Uganda  | +                  | t(2;8)      | $\mu$       | 2        | 0           | 1‡        | $\mu$    | 75              | 0   | 0  | 0    |
| JI                    | F       | 34     | Germany | +                  | t(2;8)      | $\mu\delta$ | 75       | 0           | 75        | $\mu$    | 37              | 0   | 37 | 70   |
| JBL2                  | M       | 29     | Japan   | +                  | t(2;8)      | $\mu$       | 100      | 0           | 0         | $\mu$    | 69              | 0   | 0  | 95   |
| Fresh material        |         |        |         |                    |             |             |          |             |           |          |                 |     |    |      |
| Pil                   | M       | 10     | France  | ND                 | t(8;14)     | $\mu$       | 100      | 0           | 100       | $\mu$    | 0               | 0   | 0  | ND   |
| Saa                   | M       | 12     | Tunisia | ND                 | t(8;14)     | $\mu$       | 60       | 60          | 0         | ND       | ND              | ND  | ND | ND   |
| Bas                   | M       | 76     | France  | ND                 | t(8;14)     | $\mu$       | 100      | 100         | 0         | ND       | ND              | ND  | ND | ND   |
| Tis                   | M       | 08     | France  | ND                 | t(8;14)     | $\mu$       | 100      | 100         | 0         | ND       | ND              | ND  | ND | ND   |
| Jup                   | M       | 04     | France  | —                  | t(8;14)     | $\mu$       | 100      | 0           | 100       | ND       | ND              | ND  | ND | ND   |
| Les                   | F       | 29     | France  | ND                 | t(8;14)     | $\mu\delta$ | 97       | 0           | 97        | $\mu$    | 0               | 0   | 0  | 0    |
| Cal                   | M       | 08     | France  | —                  | t(8;14)     | $\gamma$    | 93       | 93          | 0         | $\mu$    | 0               | 0   | 0  | 93   |
| Bat                   | M       | 11     | France  | —                  | t(8;14)     | $\mu$       | 100      | 100         | 0         | ND       | ND              | ND  | ND | 0    |
| Vib                   | M       | 25     | France  | ND                 | t(8;14)     | $\mu$       | 70       | 0           | 70        | ND       | ND              | ND  | ND | 0    |
| Erm                   | M       | 65     | France  | —                  | t(8;22)     | $\gamma$    | 100      | 100         | 0         | ND       | ND              | ND  | ND | 0    |

The 19 lymphoma lines were established spontaneously without addition of EBV. One line, IARC BL 14, was derived from a Burkitt-type leukaemia; three lines, Daudi, Raji and Maku, were provided by G. Klein (Stockholm); and three others, CHEV, JI and JBL2, were provided by M. Fellous (Paris), G. Bornkamm (Freiburg) and I. Miyoshi (Okayama), respectively. The MAKU line was once proposed to be one of the occasional lymphoblastoid lines established from BL tumours<sup>2</sup>. However, as no definite characterization has been made (K. Nilsson and G. Klein, personal communication), and since this line is monoclonal and possesses a characteristic BL translocation, it probably represents a true proliferation of the malignant cells. Fresh material corresponds to BL-type leukaemias, except for case Les, which is a lymphoma. The presence of EBV markers within the malignant cells was assessed by detecting the EBV nuclear antigen using anti-complement immunofluorescence. Immunoglobulin expression was evaluated using specific antisera prepared by J.L.P.; these reagents were carefully checked for absence of heterophile activity and for their specificity as discussed previously<sup>13</sup>. ND, not determined.

\* Values indicate the percentage of positive cells.

† Complex translocation, t(2;8;9), as reported elsewhere<sup>19</sup>.

‡ Despite the low percentage, this finding is of clear significance. Indeed, a large number of cells were studied, and many  $\kappa$ -positive cells were observed in the complete absence of any cell staining for  $\lambda$ .

intracytoplasmic  $\mu$  chains) to a mature B cell phenotype (with SIg, including surface IgD or IgG, and IgG Fc receptors). Translocation t(2;8) may be associated with a relatively immature phenotype or, at least, with poor light chain expression, as pre-B (IARC/LY66, JBL2) or transitional pre-B/B (IARC/LY91) phenotypes were found in 3 out of 5 cases with the t(2;8) translocation, but in only 1 out of 8 with t(8;22) and in 1 out of 19 with t(8;14).

Marked differences in the type of light chain were observed, depending on the type of translocation. For the 18 cases of t(8;14), the surface (or, when detected, intracytoplasmic) light chains were of the  $\kappa$  type in 9 cases and of the  $\lambda$  type in 8; the cells from the remaining lines had no detectable light chains. However, in 3 of 5 cases of t(2;8) in which light-chain expression was detectable, only the  $\kappa$  type was observed; conversely, in 5 of 6 cases of t(8;22),  $\lambda$  chains were detected, the remaining case corresponding to the pre-B phenotype. A summary of the results of this investigation and those of two recent reports<sup>14,15</sup> on variant translocations in BL (Table 2) shows that the three BL cases with t(2;8) that expressed immunoglobulin light chains all expressed  $\kappa$ , whereas all seven cases with t(8;22) expressed  $\lambda$ . It can be concluded, therefore, that when a light chain is expressed in BL cells its type is strictly related, in cases of variant translocation, to the chromosomes involved:  $\kappa$  in t(2;8) translocation and  $\lambda$  in t(8;22) translocation. In contrast, the light chain types were equally distributed among cases with t(8;14).

The chromosome breakpoints are thought to be constant in BL cells and are located, respectively, on 8q24 and 2p12, 14q32 and 22q11. Structural genes encoding immunoglobulin heavy

chains and  $\lambda$  chains have been assigned to chromosomes 14 and 22, respectively, without regional localization. On chromosome 2,  $\kappa$ -chain genes are located between 2pcen and 2p13 (ref. 9)—note that the breakpoint on chromosome 2 (2p12) is included within these limits. This, together with the present results, suggests that BL translocations may involve the chromosome regions near immunoglobulin loci. Therefore, precise localizations on human immunoglobulin genes may correspond to BL breakpoints, that is, 14q32, 2p12 and 22q11, for heavy chain and  $\kappa$  and  $\lambda$  genes, respectively.

Recent investigations on virally induced animal leukaemia<sup>16</sup> strongly suggest that transformation results from transposition of DNA segments such as promoter sequences. The question arises as to whether the DNA rearrangements associated with reciprocal chromosomal translocations are analogous to such DNA transpositions. The role of chromosome 8, consistently involved in BL, remains to be elucidated. It is possible that it carries on its long arm (on q24) DNA sequences involved in

Table 2 Immunoglobulin light chain types in BL

| Type of translocation | No. of cases | Light chain type |           |                |
|-----------------------|--------------|------------------|-----------|----------------|
|                       |              | $\kappa$         | $\lambda$ | Not detectable |
| t(8;14)               | 18*          | 9                | 8         | 1              |
| t(2;8)                | 5*           | 3                | 0         | 2              |
| t(8;22)               | 6*           | 0                | 5         | 1              |
|                       | 2†           | 0                | 2         | —              |

\* Results of the present study.

† Refs 14, 15.



the transformation process which are activated following translocation. The t(6;15) translocation observed in murine plasmacytomas<sup>17</sup> is another example in which malignancy, restricted immunoglobulin expression and chromosomal changes are associated, suggesting, as in BL, a position effect. In this context, the analogous involvement in lymphoid malignancies of human chromosome 8 and murine chromosome 15, recently emphasized by Klein<sup>18</sup>, is of major interest.

The correlation between immunoglobulin expression and chromosomal rearrangements reported here strongly suggests that following chromosomal translocation, an immunoglobulin gene and a putative oncogene become adjacent and are simultaneously activated under the influence of the same transcriptional element. It remains, however, to ascertain whether the immunoglobulin gene-carrying chromosome involved in the translocation is that which directs immunoglobulin

polypeptide synthesis in the malignant cell.

Elucidation of the role of non-random translocations in cancer can now be envisaged using Burkitt's lymphoma, a unique human model for which numerous well characterized lymphoma lines are available. In view of the fact that physiological DNA rearrangements occur in immunoglobulin gene clusters, a molecular approach can now be used to analyse the precise organization of DNA in the neighbourhood of the chromosome breakpoints in BL translocations, to evaluate the importance of such translocations in cancer causation.

We thank Jean-Claude Brouet and Larry B. Vogler (Hôpital St Louis) and Michèle Vuillaume (IARC) for their participation, and George Klein for his continued interest and communication of unpublished results. This work was supported in part (J.L.P.) by grant 80.7.220 from the Délégation Générale de la Recherche Scientifique et Technique.

Received 23 December 1981; accepted 7 May 1982.

1. Preud'homme, J. L., Brouet, J. C., Danon, F., Flandrin, G. & Schaison, G. *J. natn. Cancer Inst.* **66**, 261-264 (1981).
2. Zoch, L., Haglund, U., Nilsson, K. & Klein, G. *Int. J. Cancer* **17**, 47-56 (1976).
3. Van den Berghe, H. *et al. Cancer Genet. Cytogenet.* **1**, 9-14 (1979).
4. Miyoshi, I., Hiraki, S., Kimura, I., Miyamoto, K. & Sato, J. *Experientia* **35**, 742 (1979).
5. Berger, R. *et al. Hum. Genet.* **53**, 111-112 (1979).
6. Bernheim, A., Berger, R. & Lenoir, G. *C. r. hebdo. Séanc. Acad. Sci., Paris* **291**, 237-239 (1980).
7. Croce, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 3416-3419 (1979).
8. Hobart, M. J. *et al. Ann. hum. Genet.* **45**, 331-335 (1981).
9. Malcolm, S. *et al. 6th Int. Workshop on Human Gene Mapping*, Oslo, 29 June-3 July (National Foundation-March of Dimes, New York, 1981).

10. McBride, O. W., Swan, D., Leder, P., Hieter, P. & Hollis, G. *6th Int. Workshop on Human Gene Mapping*, Oslo, 29 June-3 July (National Foundation-March of Dimes, 1981).
11. Erikson, J., Martinis, J. & Croce, C. M. *Nature* **294**, 173-175 (1981).
12. I.C.S.N. *An Int. System for Human Cytogenetics Nomenclature. Birth Defects Original Article Ser.*, Vol. 14, 8 (1978).
13. Preud'homme, J. L. & Labaume, S. in *In vitro Methods in Cell-mediated and Tumor Immunity* (eds Bloom, B. R. & David, J. J.) 155-169 (Academic, New York, 1976).
14. Miyoshi, I. *et al. New Engl. J. Med.* **304**, 734 (1981).
15. Tanzer, J., Frocraim, C., Alcalay, D. & Desmaret, M. C. *Nouv. Revue fr. Hemat.* **22**, Suppl., 107 (1980).
16. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-480 (1981).
17. Wiener, F., Babonits, M., Spira, J., Klein, G. & Potter, M. *Somatic Cell Genet.* **6**, 731-738 (1980).
18. Klein, G. *Nature* **294**, 313-318 (1981).
19. Philip, T. *et al. Int. J. Cancer* **28**, 417-420 (1981).

## Cells regulate their mitogenic response to thrombin through release of protease nexin

David A. Low\*, Randy W. Scott†, Joffre B. Baker† & Dennis D. Cunningham\*

\* Department of Microbiology, California College of Medicine, University of California, Irvine, California 92717, USA

† Department of Biochemistry, University of Kansas, Lawrence, Kansas 66044, USA

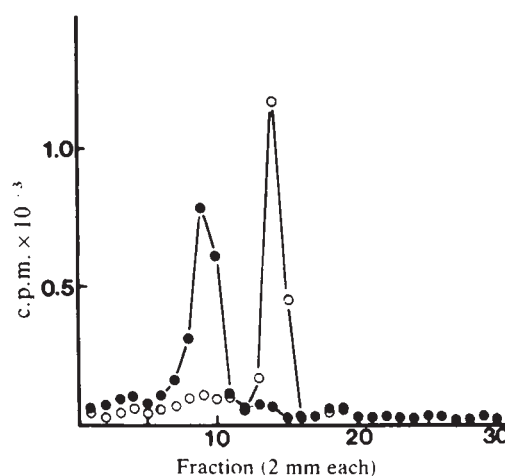
We previously reported that human and mouse fibroblast-like cells release into their growth medium a protein that we termed protease nexin. Protease nexin forms a covalent acyl linkage with thrombin and certain other serine proteases via the protease active site<sup>1</sup> and mediates their binding, internalization and degradation by cells<sup>2</sup>. Binding of thrombin-protease nexin to cells is mediated by the protease nexin portion of the complex to a high-affinity cellular binding site<sup>2</sup>. As thrombin is a potent mitogen for a variety of fibroblast-like cells in culture<sup>3-5</sup>, we examined whether protease nexin itself regulates thrombin-stimulated cell division. Recently, we showed that heparin virtually blocked the binding of thrombin-protease nexin complexes to both mouse and human cells without affecting the ability of these cells to respond to thrombin<sup>6</sup>. Thus, protease nexin does not appear to be a positive modulator in thrombin-induced cell division. Here, we show that protease nexin negatively regulates the mitogenic response of cells in culture to thrombin.

Our previous results indicated that mouse cell-conditioned medium contains substantial amounts of protease nexin<sup>1,2</sup>. Figure 1 shows the formation of <sup>125</sup>I-labelled thrombin-protease nexin complexes after addition of <sup>125</sup>I-thrombin (0.14 nM) with or without unlabelled thrombin (280 nM) to mouse cells in conditioned medium. These results show the linkage of thrombin to protease nexin in conditioned medium and demonstrate that this linkage is stable in the conditions of our assay.

Figure 2a shows a comparison of the amount of thrombin in complex with protease nexin in culture medium following

addition of increasing amounts of <sup>125</sup>I-thrombin to mouse cells in fresh or conditioned medium. In conditioned medium at <sup>125</sup>I-thrombin concentrations up to 1.4 nM, over 90% of <sup>125</sup>I-thrombin was in complex with protease nexin, compared with only 10-20% in fresh medium. Additionally, a comparison of the amount of <sup>125</sup>I-thrombin-protease nexin formed in fresh medium versus medium conditioned for 6 h shows that accumulation of protease nexin in the medium was gradual (Fig. 2a).

Cellular binding of unlinked thrombin was significantly reduced using cells in conditioned versus fresh medium (Fig. 2b). Comparison of parts a and b of Fig. 2 shows that this reduction paralleled the reduction in the amount of unlinked <sup>125</sup>I-thrombin in conditioned versus fresh medium. With increasing <sup>125</sup>I-thrombin concentration, cellular binding of



**Fig. 1** Linkage of thrombin to protease nexin. Secondary mouse embryo cells ( $9.4 \times 10^4$  cells  $\text{cm}^{-2}$  in 35-mm dishes) were cultured as described elsewhere<sup>5</sup> and washed and incubated in serum-free Dulbecco-Vogt modified Eagle's medium containing 0.1% ovalbumin (DMEM-OV). After 5 days, human <sup>125</sup>I-thrombin ( $9.3 \times 10^8$  c.p.m.  $\text{nmol}^{-1}$ , 0.14 nM) with and without addition of unlabelled thrombin (280 nM) was added without changing the medium. After 1 h at 37 °C, medium aliquots were analysed on Weber-Oshorne SDS-polyacrylamide gels<sup>11</sup>. Samples were analysed in reducing conditions but were not heated before electrophoresis. Radioactivity in each fraction was determined using a Beckman γ counter. ●, No addition; ○, addition of unlabelled thrombin.

unlinked  $^{125}\text{I}$ -thrombin in conditioned and fresh medium approached the same value. Notably, this value was not affected by the presence of ninefold more cell-bound  $^{125}\text{I}$ -thrombin-protease nexin complexes in conditioned versus fresh medium (Fig. 2b). These results support our previous finding that thrombin and thrombin-protease nexin bind to different cellular sites. We recently showed that the use of  $^{125}\text{I}$ -thrombin as a probe to determine the amount of unlinked thrombin bound to mouse cells caused about a twofold underestimation of the actual amount bound<sup>7</sup>. Here we used  $^{125}\text{I}$ -thrombin for comparative purposes only.  $^{125}\text{I}$ -thrombin is an accurate probe for measuring the formation of thrombin-protease nexin complexes<sup>7</sup>.

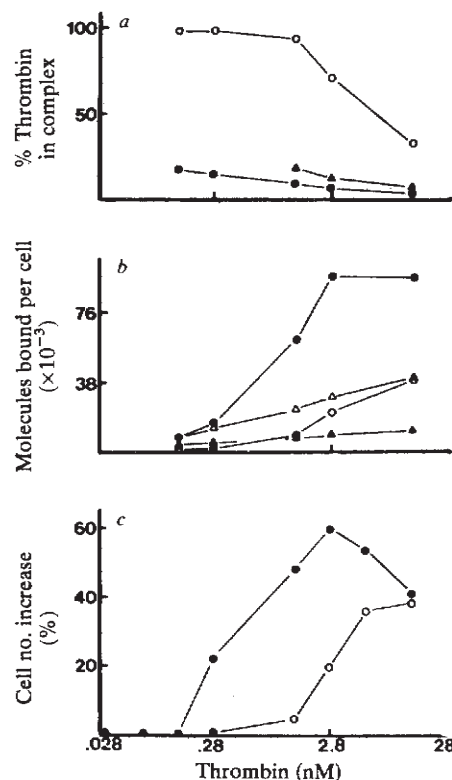
Figure 2c compares the ability of thrombin to induce mouse cell division in conditioned versus fresh medium, and shows that there was an approximately 10-fold decrease in the sensitivity of cells to thrombin in conditioned medium. Importantly, comparison of Fig. 2a and c shows that linkage of thrombin by protease nexin accounted for this decrease in mitogenic sensitivity. These results strongly indicate that mouse cell-released protease nexin negatively regulates the mitogenic response of these cells to thrombin through the formation of thrombin-protease nexin complexes.

The role of protease nexin in regulating the mitogenic action of thrombin was analysed further using a selected subclone (IIC9) of the CHEF 18 Chinese hamster embryo fibroblast cell line<sup>8</sup>, which is responsive to thrombin at extremely low doses, as evidenced by cell number (not shown) and  $^3\text{H}$ -thymidine incorporation (Fig. 3a). Addition of  $^{125}\text{I}$ -thrombin to IIC9 cell-conditioned medium resulted in the appearance of a doublet on Weber-Osborne gels of about the same size as the  $^{125}\text{I}$ -thrombin-protease nexin complex formed in mouse cell-conditioned medium (not shown). Because of the high sensitivity of IIC9 cells to the mitogenic action of thrombin, we could not analyse the amount of  $^{125}\text{I}$ -thrombin in complex with IIC9 protease nexin over a mitogenic range. However, addition of  $^{125}\text{I}$ -thrombin (28 pM) to IIC9 cells in conditioned medium resulted in linkage of ~60% of the  $^{125}\text{I}$ -thrombin to IIC9 protease nexin. Thus, IIC9 cells produced significant levels of protease nexin, although less than the amount produced by mouse cells (Figs 1, 2a).

Figure 3a shows that conditioning of the medium by IIC9 cells had two effects on thrombin-mediated mitogenesis. At low doses of thrombin (0.28–1.4 pM), cells in conditioned medium were about fivefold less sensitive to thrombin action than cells in fresh medium, whereas at higher doses of thrombin, cells responded up to twofold better in conditioned medium. Figure 3b shows that conditioned medium increased the mitogenic response of IIC9 cells to epidermal growth factor over a wide concentration range for this hormone. Thus, IIC9-conditioned medium contains a factor which decreases the sensitivity of cells to thrombin but not to epidermal growth factor. By analogy with the results presented for mouse cells (Fig. 2), it seems likely that this factor is IIC9 protease nexin.

Figure 3a shows the effect of a heparin-Sepharose affinity-purified preparation of protease nexin on the response of IIC9 cells to thrombin. Protease nexin constituted ~10% of the total protein in this preparation (~300-fold purification) and was the only thrombin inhibitor detected (R.W.S. and J.B.B., unpublished data). At a dilution of  $2.5 \times 10^{-3}$  (0.14 nM, as determined by linkage to saturating amounts of  $^{125}\text{I}$ -thrombin), this preparation decreased the sensitivity of IIC9 cells to thrombin by about 10-fold. However, this same amount of partially purified protease nexin did not affect the response of IIC9 cells to epidermal growth factor (Fig. 3b) or fetal calf serum (not shown). These results indicate that protease nexin specifically inhibited the mitogenic response of IIC9 cells to thrombin.

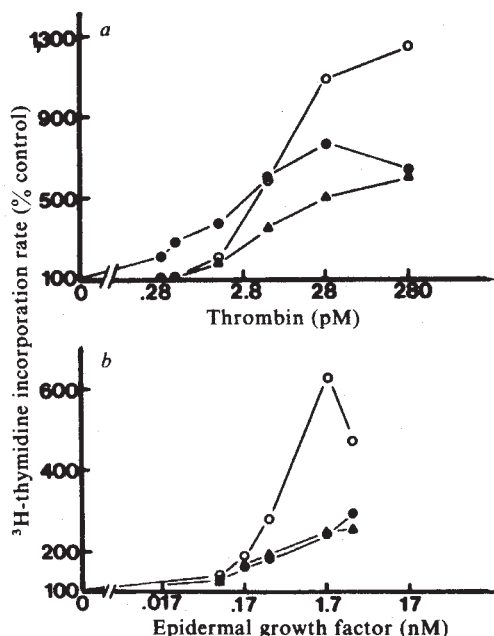
Our recent results indicate that binding of thrombin to the cell-surface receptor for unlinked thrombin on mouse and IIC9 cells is not necessary for thrombin-stimulated cell division (D.A.L. and D.D.C., unpublished data). Rather, it seems likely that thrombin initiates cell division by cleavage<sup>9</sup> of a cell-surface protein(s)<sup>10</sup>. Because linkage of thrombin to protease nexin



**Fig. 2** The effects of protease nexin and conditioned medium on the cellular binding and mitogenic activity of thrombin in cultures of mouse cells. **a**, Increasing amounts of  $^{125}\text{I}$ -thrombin were added to mouse cell cultures as described in Fig. 1 legend except that after a 5-day incubation in DMEM-OV, medium from some cultures was removed and replaced with fresh DMEM-OV. Formation of  $^{125}\text{I}$ -thrombin-protease nexin complexes in the medium was analysed as described in Fig. 1 legend. ○,  $^{125}\text{I}$ -thrombin-protease nexin formation after addition of  $^{125}\text{I}$ -thrombin to cells in medium conditioned for 5 days; ▲, medium conditioned for 6 h; ●, fresh medium. **b**,  $^{125}\text{I}$ -thrombin, in the presence and absence of 280 nM unlabelled thrombin, was incubated with mouse cell cultures as described in Fig. 1 legend. After six rinses with Dulbecco's phosphate-buffered saline, cells were dissolved in 200  $\mu\text{l}$  of bis-Tris solubilization buffer as described previously<sup>2</sup> except that samples were not heated. SDS-soluble proteins were analysed using 1 mm-thick polyacrylamide slab gels consisting of a 3% acrylamide/bis-Tris stacking gel and a 7.5–15% acrylamide/mono-Tris separating gel as described<sup>14</sup>. Autoradiograms of slab gels were prepared as described previously<sup>15</sup> and bands containing  $^{125}\text{I}$ -thrombin and  $^{125}\text{I}$ -thrombin-protease nexin were excised and quantified as described in Fig. 1 legend. The amount of radioactivity in each sample was normalized based on the fraction: [total radioactivity recovered from gel/total radioactivity in sample], which was about 0.8. The amount of nonspecific binding (binding in the presence of 280 nM thrombin) was subtracted to yield the amount of specifically bound  $^{125}\text{I}$ -thrombin and  $^{125}\text{I}$ -thrombin-protease nexin at each  $^{125}\text{I}$ -thrombin concentration. ●, Cell-bound  $^{125}\text{I}$ -thrombin-protease nexin using cells in medium conditioned for 5 days; ○, cell-bound  $^{125}\text{I}$ -thrombin using cells in medium conditioned for 5 days; ▲, cell-bound  $^{125}\text{I}$ -thrombin-protease nexin using cells in fresh medium; △, cell-bound  $^{125}\text{I}$ -thrombin using cells in fresh medium. **c**, Mouse cells were added to 35-mm dishes ( $6 \times 10^4$  cells  $\text{cm}^{-2}$ ) in DMEM containing 5% calf serum. After 6 h, cells were rinsed twice and incubated in 2 ml of serum-free DMEM-OV. After 5 days, medium from some of the dishes was removed and replaced with fresh DMEM-OV. Two days after addition of thrombin, the number of cells on each dish was determined using a Coulter particle counter. Control cultures (no addition of thrombin) in conditioned medium contained  $7.0 \times 10^4$  cells  $\text{cm}^{-2}$  whereas control cultures in fresh medium contained  $4.8 \times 10^4$  cells  $\text{cm}^{-2}$ . However, cultures in fresh medium receiving non-mitogenic levels of thrombin (0.028–0.14 nM) contained the same number of cells as control cultures in conditioned medium. Apparently, non-mitogenic levels of thrombin were able to maintain cell viability and/or adhesion to the dish in the absence of cell-released conditioning factors. A similar phenomenon has recently been reported using insulin and epidermal growth factor<sup>16</sup>. Therefore, the same control culture cell number ( $7.0 \times 10^4$  cells  $\text{cm}^{-2}$ ) was used for both conditioned and fresh media. ○, Medium conditioned for 5 days; ●, fresh medium.

results in loss of thrombin proteolytic activity<sup>6,11</sup>, it seems likely that protease nexin regulates the mitogenic activity of thrombin by regulating thrombin proteolysis. Induction of the release of protease nexin from bovine aortic endothelial cells<sup>12</sup> and human fibroblasts by thrombin (D.L. Eaton and J.B.B., unpublished





**Fig. 3** The effects of conditioned medium and a fraction containing partially purified protease nexin on the mitogenic response of IIC9 cells to thrombin and epidermal growth factor. IIC9 cells were cultured chiefly as described for CHEF 18 cells<sup>17</sup> using MEM- $\alpha$ /Ham's F12 medium (1:1, v/v) containing 10% fetal calf serum and HEPES (10 mM). Cells in this medium containing 5% fetal calf serum were seeded into 35-mm dishes ( $10^4$  cells  $\text{cm}^{-2}$ ) and incubated for 5–6 days, by which time all cultures had reached confluence at  $6\text{--}7 \times 10^4$  cells  $\text{cm}^{-2}$ . Cultures were then rinsed three times with 2 ml of serum-free MEM- $\alpha$  without ribonucleosides and deoxyribonucleosides/Ham's F12 medium (1:1, v/v) containing human transferrin ( $5 \mu\text{g ml}^{-1}$ ; Sigma), bovine insulin ( $10 \mu\text{g ml}^{-1}$ ) (Sigma), fibroblast growth factor ( $10 \text{ ng ml}^{-1}$ ; a gift from D. Gospodarowicz) and ovalbumin (0.1%), and 2 ml of this medium (3F medium) was added to each dish. The indicated amounts of thrombin and epidermal growth factor were either added immediately or after 1 day of incubation. 18 h after hormone addition medium was withdrawn and replaced with 1 ml of fresh 3F medium containing  $^3\text{H}$ -thymidine ( $1 \mu\text{Ci ml}^{-1}$ ). After 2 h of incubation, cell cultures were processed for  $^3\text{H}$ -thymidine incorporation as described previously<sup>18</sup>. Values are expressed relative to control dishes which did not receive hormone. About 15% of cells on control dishes contained labelled nuclei after a 23 h exposure to  $^3\text{H}$ -thymidine (added 6 h after thrombin addition), whereas maximal levels of thrombin resulted in ~65% labelled nuclei. Additionally, 3F medium which had first been incubated with mitogenic amounts of thrombin and then with hirudin to inactivate thrombin was not mitogenically stimulatory (not shown). These results showed that thrombin must interact with IIC9 cells in order to stimulate them mitogenically. *a*, Thrombin and *b*, epidermal growth factor addition to cells in  $\bullet$ , fresh medium;  $\blacktriangle$ , fresh medium with addition of 5  $\mu\text{l}$  per dish of a fraction containing partially purified protease nexin, derived from human cell-conditioned medium;  $\circ$ , conditioned medium.

data) could provide a feedback mechanism for both mitogenic autoregulation and regulation of the mitogenic response of neighbouring cells to thrombin.

This work was supported by NIH research grants CA-12306 and CA-29307, and by a grant from the American Heart Association, Kansas Affiliates. D.A.L. was supported by pre-doctoral training grant GM-07134. We thank J.W. Fenton II for thrombin, H.S. Wiley for providing IIC9 cells, D. Knauer and D. Raben for useful discussion and G. Ryan for technical assistance.

Received 5 February; accepted 19 May 1982.

1. Baker, J. B., Low, D. A., Simmer, R. L. & Cunningham, D. D. *Cell* **21**, 37–45 (1980).
2. Low, D. A., Baker, J. B., Koonce, W. C. & Cunningham, D. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2340–2344 (1981).
3. Chen, L. B. & Buchanan, J. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 131–135 (1975).
4. Pohjanpelto, P. *J. cell. Physiol.* **95**, 189–194 (1978).
5. Carney, D. H., Glenn, K. C. & Cunningham, D. D. *J. cell. Physiol.* **95**, 13–22 (1978).
6. Baker, J. B., Low, D. A., Eaton, D. L. & Cunningham, D. D. *J. cell. Physiol.* (in the press).
7. Low, D. A. & Cunningham, D. D. *J. biol. Chem.* **257**, 850–858 (1982).
8. Sager, R. & Kovac, P. E. *Somatic Cell Genet.* **4**, 375–392 (1978).
9. Glenn, K. C., Carney, D. H., Fenton, J. W. II & Cunningham, D. D. *J. biol. Chem.* **255**, 6609–6616 (1980).
10. Carney, D. H. & Cunningham, D. D. *Cell* **14**, 811–823 (1978).

11. Low, D. A., Cunningham, D. D., Scott, R. W. & Baker, J. B. in *Receptor-Mediated Binding and Internalization of Toxins and Hormones* (eds Middlebrook, J. L. & Kohn, L. D.) 259–270 (Academic, New York, 1981).
12. Savion, N., Isaacs, J. D., Gospodarowicz, D. & Shuman, M. A. *J. biol. Chem.* **256**, 4514–4519 (1981).
13. Weber, K. & Osborne, M. J. *J. biol. Chem.* **244**, 4406–4412 (1969).
14. Knauer, D. J. & Cunningham, D. D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2310–2314 (1982).
15. Baker, J. B., Simmer, R. L., Glenn, K. C. & Cunningham, D. D. *Nature* **278**, 743–745 (1979).
16. Berliner, J. A., Aharanov, A. & Pruss, R. *Expl Cell Res.* **133**, 227–235 (1981).
17. Cherington, P. V. & Pardee, A. B. *J. cell. Physiol.* **105**, 25–32 (1980).
18. Knauer, D. J. & Smith, G. L. *J. cell. Physiol.* **100**, 311–322 (1979).

## Physiological $[\text{Ca}^{2+}]_i$ level and pump-leak turnover in intact red cells measured using an incorporated Ca chelator

Virgilio L. Lew, Roger Y. Tsien\* & Cristina Miner

Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, UK

Robert M. Bookchin

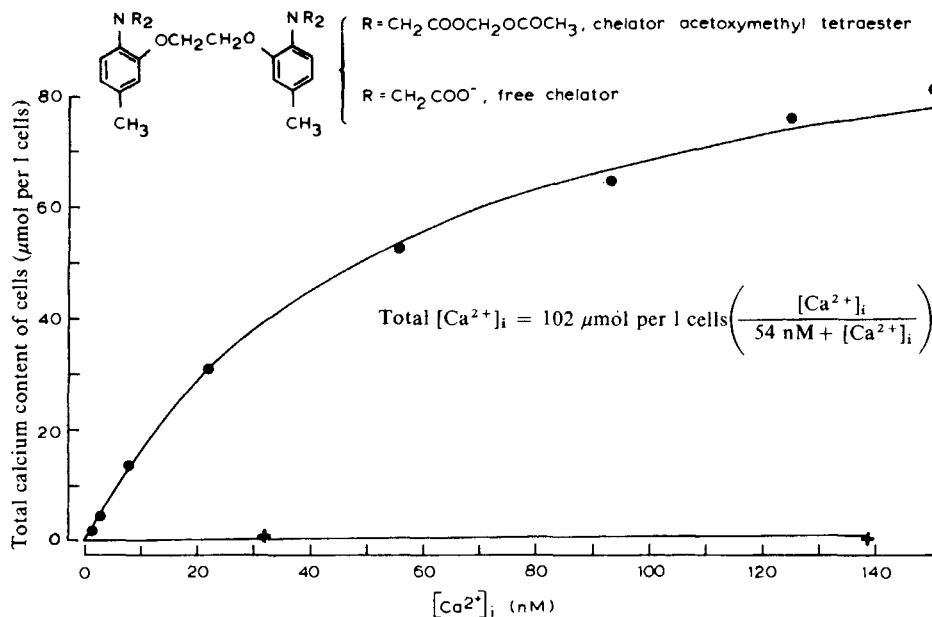
Department of Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA

The physiological actions of  $\text{Ca}^{2+}$  as a trigger and second messenger depend on the maintenance of large inward resting  $\text{Ca}^{2+}$  gradients across the cell plasma membrane. An ATP-fuelled Ca-pump, originally discovered and still best characterized in human red cells<sup>1–4</sup>, is now believed to mediate resting  $\text{Ca}^{2+}$  extrusion in most animal cells. However, even in red cells, the truly physiological pump-leak turnover rate and cytoplasmic free  $\text{Ca}^{2+}$  level are unknown. Previous estimates<sup>5,6</sup> were only very imprecise upper limits because normal intact red cells have a minute total pool of exchangeable Ca of less than  $1 \mu\text{mol l}^{-1}$  cells<sup>7</sup>; Ca fluxes could not be measured without artificially increasing that pool with ionophores<sup>8–10</sup> or disrupting the membrane to incorporate Ca buffers<sup>5</sup>. Both procedures leave the membrane considerably leakier than in intact cells. Here, we have increased the exchangeable Ca pool by non-disruptively loading a Ca-chelator into intact cells, using intracellular hydrolysis of a membrane-permeant ester<sup>11,12</sup>. The trapped chelator made the free cytoplasmic calcium concentration,  $[\text{Ca}^{2+}]_i$ , an easily defined function of directly measurable total cell Ca. We were then able to establish the physiological steady-state  $[\text{Ca}^{2+}]_i$  and pump-leak turnover rate of fresh cells suspended in their own plasma. If  $[\text{Ca}^{2+}]_i$  was lowered below the normal resting level, the Ca pump rate decreased according to the square of  $[\text{Ca}^{2+}]_i$ , and the inward Ca leak increased. The increase in leak did not develop if the cells were depleted of ATP and ADP.

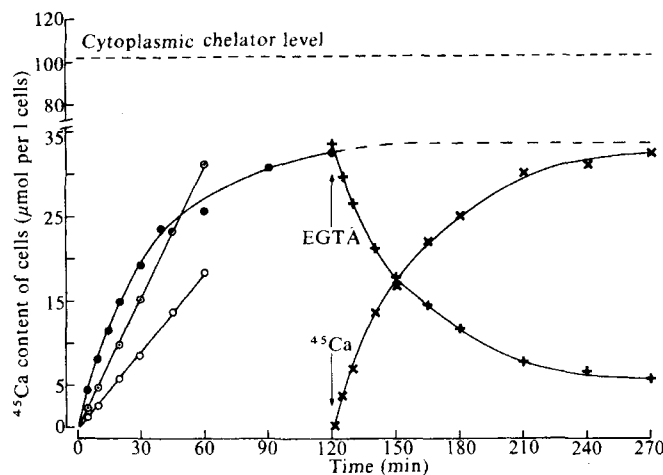
A typical experiment is illustrated in Figs 1 and 2. Fresh cells in Ca-free saline were loaded with chelator as described in Fig. 1 legend. An aliquot of these cells was analysed by equilibration with Ca-EGTA buffers and Ca ionophore to determine the *in situ* Ca-binding properties of the intracellular chelator. As detailed in Fig. 1 legend, the loaded cells behaved as if they contained  $102 \mu\text{mol per l}$  cells of a buffer with 54 nM dissociation constant for  $\text{Ca}^{2+}$ . Another aliquot of the chelator-loaded cells was put back into autologous plasma containing  $^{45}\text{Ca}$  and additional glucose. The total Ca content of these cells (Fig. 1, filled circles) increased gradually over 2 h, levelling off at  $33 \mu\text{mol l}^{-1}$  cells, a value much larger and more readily quantifiable than unloaded cells would show, yet corresponding to only 26 nM  $[\text{Ca}^{2+}]_i$  based on the calibration curve in Fig. 1. In two other experiments on cells and plasma from the same donor as in Fig. 2, the steady-state  $[\text{Ca}^{2+}]_i$  values were 11 and

\* Present address: Department of Physiology-Anatomy, University of California, Berkeley, California 94720, USA.

**Fig. 1** Determination of the intracellular chelator content and Ca-binding kinetics in intact red cells. Because of the intense pigmentation of red cells,  $[Ca^{2+}]_i$  could not be monitored from the spectral shifts of the chelator as was done in transparent cells<sup>18</sup>. The low endogenous Ca-binding capacity of red cells at low  $[Ca^{2+}]_i$  (see control curve) offers an alternative way to measure  $[Ca^{2+}]_i$  in chelator-loaded cells as virtually all the Ca will be bound to the chelator. The principle of the method was similar to that previously applied to the determination of cytoplasmic Mg buffering in intact red cells<sup>19</sup>. Chelator-loaded, ATP-depleted cells were suspended in solutions containing buffered  $^{45}Ca^{2+}$  in the nanomolar range, together with sufficient Ca ionophore A23187 to ensure that  $[Ca^{2+}]_i$  reached electrochemical equilibrium with  $[Ca^{2+}]_o$ . At equilibrium, the  $^{45}Ca$  distribution and specific activity were measured and from these the total cell Ca content at each  $[Ca^{2+}]_i$  could be obtained. An aliquot of the chelator-loaded, ATP-depleted cells, as described in Fig. 2 legend was washed twice and resuspended at a haematocrit of 8% in a medium containing (mM): NaCl, 75; KCl, 75; Tris-Cl (pH 7.6 at 37 °C), 10;  $MgCl_2$ , 0.20;  $^{45}CaCl_2$ , 0.10. The ionophore A23187 was added to this suspension (at 37 °C) from a 2 mM solution in ethanol to give a final concentration of 6  $\mu M$  in the suspension. About 10 min after ionophore addition, 1-ml portions of the cell suspension were distributed to tubes containing concentrated EGTA (40–400 mM) to give final total EGTA concentrations of 120  $\mu M$ –8 mM. After about 90 min, to allow for  $Ca^{2+}$  equilibration, triplicate 0.2-ml samples were taken at 5-min intervals from each of the conditions and processed as described in Fig. 2 legend, except for the use of high-K wash medium (containing 0.1 mM Tris-EGTA) and the presence of albumin (0.25 g%) in the first wash in order to extract the ionophore and prevent loss of  $^{45}Ca$  from the cells<sup>20</sup>. Immediately after the experiment the pH of the suspension and that of the separated, freeze-thaw-disrupted cells, was measured with a capillary Astrup-type pH meter at 37 °C.  $[Ca^{2+}]_o$  was calculated from the known EGTA and total Ca concentration at the measured pHs.  $[Ca^{2+}]_i$  was calculated from the equilibrium condition:  $[Ca^{2+}]_i = [Ca^{2+}]_o \cdot ([H^+]_i/[H^+]_o)^2$ , valid in red cells in these conditions<sup>9,9</sup>. The figure gives the relationship between measured total cell Ca and  $[Ca^{2+}]_i$  for controls without chelator (+) and for the cells used in the experiment of Fig. 2 (●). The data are well described by 1:1 binding to 102  $\mu mol l^{-1}$  cells of chelator. The *in situ* dissociation constant ( $K_D$ ) in seven different experiments was  $50 \pm 4$  nM (mean  $\pm$  s.e.m.), in adequate agreement with the 40 nM value estimated spectrophotometrically in protein-free, 100 mM KCl solutions at 22 °C (ref. 11). This shows that the chelator acts as if it were free in the cytoplasm. For any known total cell Ca content,  $[Ca^{2+}]_i$  in each experiment could be read off from the corresponding calibration curve.



**Fig. 2** Total Ca content of cells as a function of time in chelator-loaded fresh red cells suspended in their own plasma. Red cells from fresh heparinized human blood were washed three times in a solution containing (mM): NaCl, 150; KCl, 5; Tris-Cl (pH 7.6 at 37 °C), 10, and Tris-EGTA, 0.1, and resuspended in the same medium with additional glucose, 10 mM, at a haematocrit of 30%. The acetoxymethyl tetraester derivative of the chelator, structure 2b of Tsien<sup>11</sup> modified as shown in Fig. 1, was added to this suspension from a concentrated solution in dimethyl sulphoxide (DMSO, 100 mM) in amounts calculated to give  $\sim 120 \mu mol l^{-1}$  cells of incorporated free chelator if the ester had been fully hydrolysed and trapped in the cells. The addition was under vigorous vortex stirring. Stirring was maintained for about 1 min to secure a homogeneous distribution of chelator ester in all red cell membranes. The suspension was then incubated at 37 °C for 75 min during which the cells became loaded with the chelator (see Fig. 1). Four mols of formaldehyde should have been evolved per mol of chelator ester incorporated<sup>12</sup>. Control experiments with the highest DMSO and formaldehyde concentrations used, but without chelator, demonstrated no detectable effects on Ca fluxes relative to untreated controls. After chelator loading, the cells were washed in the original medium and a part of them depleted of ATP by a further 2.5 h incubation at 37 °C in the presence of (mM): inosine, 10 and iodoacetamide, 6. Ca uptake is reported in chelator-loaded, glucose-fed (●, +, ×) or ATP-depleted (○, ○) cells, suspended at 10% haematocrit and incubated at 37 °C either in autologous plasma with (●, ○, +) or without (×)  $^{45}Ca$  ( $10 \mu Ci \mu mol^{-1}$ ), or in the original wash medium with 1.3 mM  $^{45}Ca$  ( $10 \mu Ci \mu mol^{-1}$ ) instead of the Tris-EGTA (○). After 120 min at 37 °C, 4 mM NaOH-neutralized EGTA (calibrated to cause no significant change in pH) was added to condition (+) and  $^{45}Ca$ -tracer to condition (×). The  $^{45}Ca$  content of the cells was measured in 0.2-ml samples of the various cell suspensions, taken at the indicated times and added to 14 ml ice-cold original washing solution. The cells in each sample were washed twice more in this solution, lysed in 6% trichloroacetic acid, and the  $^{45}Ca$  activity measured in an aliquot of the trichloroacetic acid supernatant. The Ca content was calculated from the measured  $^{45}Ca$  content of the cells and from the specific activity of the extracellular  $^{45}Ca$ . The indicated chelator level was determined as explained in Fig. 1 legend.

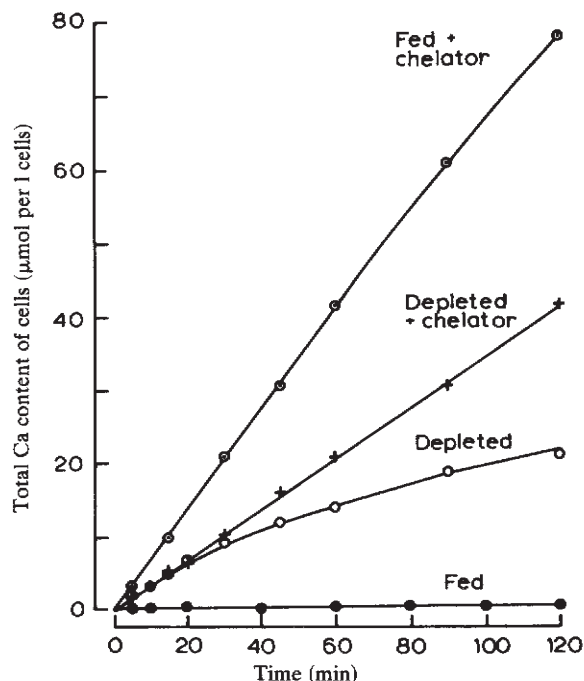


28 nM. These values should represent the physiological steady-state  $[Ca^{2+}]_i$  of red cells unless the chelator loading has modified the Ca pump or leak influx. Control experiments described below and in Fig. 2 legend have so far revealed no sign of such alterations.

The turnover rate at 120 min, very near the pump-leak steady-state, was measured (1) by adding excess EGTA over Ca and following the subsequent net Ca release and (2) by adding  $^{45}Ca$  to a parallel condition without tracer initially and following the tracer equilibration curve (see Fig. 2). In these conditions, the net Ca release represents Ca-pump-mediated efflux, as addition of excess EGTA to chelator-loaded, ATP-depleted cells inhibits further uptake but causes no measurable

release (not shown). The tracer-distribution flux measures inward leak and Ca:Ca exchange. The turnover fluxes in the experiment of Fig. 2 were 45 (efflux) and 48 (influx)  $\mu mol$  per l cells per h, respectively. The similarity between the two unidirectional fluxes at such different external  $Ca^{2+}$  concentrations (about 50 nM during net efflux and the normal  $Ca^{2+}$  plasma concentration of 1.3 mM in the tracer influx condition) suggests that this flux (45–48  $\mu mol$  per l cells per h) represents the true Ca turnover rate in physiological conditions. In the experiment of Fig. 2, alkalization in air was prevented by titration with HCl to constant pH at 7.4. In another similar plasma experiment, the pH was maintained at 7.4 under 5%  $CO_2$  atmosphere and the steady-state turnover fluxes (efflux/influx) were

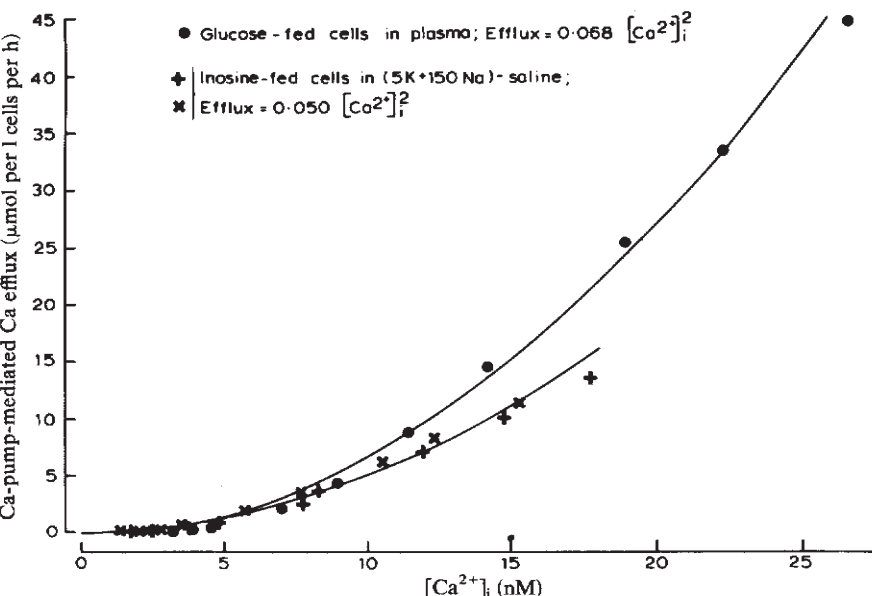




**Fig. 3** Uptake of Ca by inosine-fed or ATP-depleted red cells with and without intracellular chelator. Red cells from fresh heparinized human blood were washed and loaded with chelator as described in Fig. 2 legend, except that the solution used throughout was (mM): NaCl, 75; KCl, 75; Tris-Cl (pH 7.6 at 37 °C) and inosine, 10. The chelator-free controls (○, ●) were treated in the same way as the chelator-loaded cells (+, ×) except for the presence of chelator ester in the added DMSO. The chelator content of the cells was about 0.6 mM. The chelator incorporation yield in inosine-fed cells was only 20–25% whereas in two experiments with glucose-fed cells it was ~90%. The reason for this has not yet been investigated. Internal  $P_i$ , which is much reduced in inosine-fed human red cells<sup>15</sup>, may be required by the membrane esterases which hydrolyse the chelator ester, releasing free chelator inside the cells. ATP depletion was performed after chelator loading by incubation for 2.5 h at 37 °C in the presence of inosine, 10 mM and iodoacetamide, 6 mM (○, +). The final incubation medium contained  $^{45}\text{CaCl}_2$ , 1 mM.

30/32  $\mu\text{mol per 1 cell per h}$  ( $[\text{Ca}^{2+}]_i = 28 \text{ nM}$ ).

The results in Fig. 2 contain some unexpected findings. The turnover flux is considerably higher than earlier estimates of 10–20  $\mu\text{mol per 1 cell per h}$  for ATP-depleted cells in low-K saline<sup>13,14</sup>. This difference is found to be due partly to the substitution of saline for plasma and partly to ATP depletion, but not to chelator incorporation *per se*. Thus, Fig. 2 shows that



**Fig. 4** Ca-pump mediated Ca efflux as a function of  $[\text{Ca}^{2+}]_i$  in the sub-physiological  $\text{Ca}^{2+}$  concentration range. The slope of the net Ca-extrusion curve of Fig. 2 after addition of excess EGTA (+) is plotted here against  $[\text{Ca}^{2+}]_i$  as read off from the chelator calibration curve of Fig. 1 (●). The results of two similar experiments (+, ×) but with cells suspended in (mM): NaCl, 150; KCl, 5; Tris-Cl (pH 7.6 at 37 °C), 10;  $^{45}\text{CaCl}_2$ , 1.3 and inosine, 10, and with pump-leak steady-state  $[\text{Ca}^{2+}]_i$  levels of 15 (×) and 17 (+) nM, are also included. The lines correspond to the equation given in the figure. The equations are of the form  $\phi = A [\text{Ca}^{2+}]_i^2$ , where  $A$  is in units of  $\mu\text{mol per 1 cell per h per nM}^2$ ,  $\phi$  is in  $\mu\text{mol per 1 cell per h}$  and  $[\text{Ca}^{2+}]_i$  is in nM. In five different experiments in plasma and in saline, with blood from two different donors, the value of  $A$  was in the 0.050–0.140  $\mu\text{mol per 1 cell per h per nM}^2$  range. Interpretation of  $A$  in terms of  $\phi_{\text{max}}/K_1K_2$  (see text) gives  $\phi_{\text{max}}$  values around 10–100  $\mu\text{mol per 1 cell per h}$  if the  $K_1K_2$  product is in the 0.5–4  $\mu\text{M}^2$  range as previously estimated in intact red cells<sup>8,10</sup>.

Ca influx into ATP-depleted cells in their own plasma amounted to 31  $\mu\text{mol per 1 cell per h}$ , whereas equivalent cells in saline gained only 17  $\mu\text{mol per 1 cell per h}$ . These uptake rates remain constant and represent pure influx because ATP depletion has stopped the pump. Investigations are underway to determine why the Ca permeability of red cells should be higher in plasma than in saline, but direct effects of chelator loading can be ruled out. First, the Ca uptake by chelator-loaded cells suspended in buffered saline fell within the expected range<sup>14</sup>. Second, the Ca uptake in high-K media, where Ca uptake is lowest<sup>13</sup> and small effects are more likely to be exposed, was initially identical in chelator-free cells and in cells loaded with ~600  $\mu\text{mol l}^{-1}$  cells of chelator (Fig. 3). Hence, neither the chelator nor the chelator-loading procedure seem to have had any harmful effects on the Ca permeability of the red cell membrane. The results in Fig. 3 also confirm that the levelling off of the Ca-uptake curve in chelator-free ATP-depleted cells is due to self-limiting effects of internal  $\text{Ca}^{2+}$  on its own permeability, as suggested previously.

Figure 2 also illustrates the dependence of Ca influx on other unexpected influences,  $[\text{Ca}^{2+}]_i$  and ATP/ADP content. At  $[\text{Ca}^{2+}]_i < 5 \text{ nM}$ , influx into normally fed cells (solid circles in Fig. 2 near time = 0) was 56  $\mu\text{mol per 1 cell per h}$ , noticeably faster than the turnover influx at the steady-state  $[\text{Ca}^{2+}]_i$  level of 26 nM. Cells depleted of ATP and ADP (dotted circles) gained only 31  $\mu\text{mol per 1 cell per h}$ , a rate independent of  $[\text{Ca}^{2+}]_i$  in the sub-physiological concentration range. Similar results were obtained in two other experiments in plasma and also in saline, as seen in Fig. 3 by comparing dotted circles and crosses. Therefore, the unidirectional Ca influx has a metabolism-dependent component which becomes partly inhibited as  $[\text{Ca}^{2+}]_i$  increases towards its pump-leak steady-state level. It is tempting to interpret this flux as representing Ca-pump reversal when the inward  $\text{Ca}^{2+}$  gradient becomes steeper than normal. Inosine, however, which reduces the internal  $P_i$  concentration (also in chelator-loaded cells) and blocks Na-pump reversal most effectively<sup>15</sup>, had no effect on the metabolism-dependent Ca influx (Fig. 3). Hence, if the metabolism-dependent influx represents Ca-pump reversal, it is unlikely that it is coupled to ATP synthesis. Whatever its mechanism, the extra Ca influx which develops at subnormal  $[\text{Ca}^{2+}]_i$  may help tighten the regulation of  $[\text{Ca}^{2+}]_i$  by the plasma membrane.

Figure 2 also shows that the dependence of the pump rate,  $\phi$ , on  $[\text{Ca}^{2+}]_i$  can be deduced from the net efflux curve (marked by crosses). This dependence is plotted in Fig. 4 and is well described by the equation  $\phi = A [\text{Ca}^{2+}]_i^2$ . The exponent of 2 suggests that the pump is activated by the internal binding of two  $\text{Ca}^{2+}$  ions. As the curve remains far below saturation over

the entire range explored, all we can conclude about the  $\text{Ca}^{2+}$ -dissociation constants  $K_1$  and  $K_2$  of the activating sites is that  $K_1$  and  $K_2$  should be considerably greater than 26 nM and that the ratio of the eventual saturating flux to  $K_1K_2$  should equal  $A$ . Moreover, the apparent activation by two  $\text{Ca}^{2+}$  ions does not imply that both must be transported. Indeed, the free energy available from hydrolysis of 1 mol of ATP either barely equals or falls short of that required to pump 2 mol of  $\text{Ca}^{2+}$  from the cytoplasm at 26 nM and  $-9$  mV to the outside at 1.3 mM, depending on the precise values taken for free ATP, ADP and  $\text{P}_i$  in red cells<sup>16,17</sup>. However, the possibility that the flattening of the Ca-efflux curve at low  $[\text{Ca}^{2+}]_i$  (Fig. 4) is due to minute residual  $^{45}\text{Ca}$ -influx components rather than to parabolic pump activation cannot be ruled out.

We thank Mrs J. Gray for technical assistance, T. J. Rink, R. Macey and I. M. Glynn for reading the manuscript and the MRC, SRC, NIH and MEC (Spain) for funds.

Received 2 April; accepted 27 May 1982.

- Schatzmann, H. J. *Experientia* **22**, 364–368 (1966).
- Scharff, O. & Foder, B. *Biochim. biophys. Acta* **509**, 67–77 (1978).
- Muallem, S. & Karlish, S. J. D. *Biochim. biophys. Acta* **647**, 73–86 (1981).
- Niggli, V., Adunyah, E. S., Penniston, J. T. & Carafoli, E. *J. biol. Chem.* **256**, 395–401 (1981).
- Schatzmann, H. J. *J. Physiol., Lond.* **235**, 551–569 (1973).
- Roufogalis, B. D. *Can. J. Physiol. Pharmacol.* **57**, 1331–1349 (1979).
- Bookchin, R. M. & Lew, V. L. *Nature* **284**, 561–563 (1980).
- Ferreira, H. G. & Lew, V. L. *Nature* **259**, 47–49 (1976).
- Lew, V. L. & Brown, A. M. in *Detection and Measurement of Free Ca in Cells* (eds Ashley, C. C. & Campbell, A. K.) 423–432 (Elsevier, Amsterdam, 1979).
- Lew, V. L., Bookchin, R. M., Brown, A. M. & Ferreira, H. G. *Alfred Benzon Symp.* **14**, 196–207 (1980).
- Tsien, R. Y. *Biochemistry* **19**, 2396–2404 (1980).
- Tsien, R. Y. *Nature* **290**, 527–528 (1981).
- Lew, V. L. in *Comparative Biochemistry and Physiology of Transport* (eds Bolis, L., Bloch, K., Luria, S. E. & Lynen, F.) (North-Holland, Amsterdam, 1974).
- Ferreira, H. G. & Lew, V. L. in *Membrane Transport in Red Cells* (eds Ellory, J. C. & Lew, V. L.) 53–91 (Academic, New York, 1977).
- Glynn, I. M., Lew, V. L. & Lüthi, U. *J. Physiol., Lond.* **207**, 371–391 (1970).
- Pennell, R. N. in *The Red Blood Cell* 2nd edn (ed. Surgeur, D. MacN.) 93–416 (Academic, New York, 1974).
- Gupta, R. K., Benovic, J. L. & Rose, Z. B. *J. biol. Chem.* **253**, 6172–6176 (1978).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68–71 (1982).
- Flatman, P. & Lew, V. L. *J. Physiol., Lond.* **305**, 13–30 (1980).
- Simonsen, L. O. *J. Physiol. Lond.* **313**, 34P–35P (1980).

## Epifluorescent microscopic evidence for maternal inheritance of chloroplast DNA

Tsuneyoshi Kuroiwa, Shigeyuki Kawano & Soryu Nishibayashi

Division of Cell Mechanism, Department of Cell Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaijicho, Okazaki, 444 Japan

Chubun Sato

Department of Biology, Kagawa Medical School, Mikicho, Kidagun, 761-07 Japan

Maternal inheritance of chloroplast genes occurs in the isogamous green alga *Chlamydomonas reinhardtii*<sup>1,2</sup>. It has been shown<sup>3–5</sup> using biochemical techniques that the chloroplast DNA of male origin is preferentially lost by 6 h after mating. DNAs in the chloroplast are organized by proteins into about 10 chloroplast nucleoids<sup>6,7</sup>. Therefore, if chloroplast DNA in zygotes is preferentially destroyed, the disappearance of chloroplast nucleoids from male gametes should be observable during zygote formation by high resolution epifluorescent microscopy. Here we present the first fluorescent microscopic evidence that in *C. reinhardtii*, about eight chloroplast nucleoids from the male parent disappear during the first 40–50 min after mating, while those from the female parent persist and finally fuse together to form one large chloroplast nucleoid.

Gametes of both mating types ( $137c\text{mt}^+$  and  $\text{mt}^-$ ) of *C. reinhardtii* were prepared and mixed to allow mating. To distinguish the  $\text{mt}^+$  morphologically from the  $\text{mt}^-$  gametes, small  $\text{mt}^+$  gametes obtained by centrifugation at 180g for 2 min and  $\text{mt}^-$  gametes of a mutant (SF 60, Sato culture collection) with short flagella ( $\sim 3\ \mu\text{m}$  long) were used. The chloroplast nucleoids in the living gametes and zygotes were stained with a DNA-specific fluorochrome, 4'-6-diamidino-2-phenylindole (DAPI), dissolved in distilled water at  $1\ \mu\text{g}\ \text{ml}^{-1}$ , and observed with an Olympus BH-RFK epifluorescent microscope as described previously<sup>8</sup>. The fluorescent microscope was equipped with a high-pressure mercury vapour lamp (HBO; 100 W), a 334 nm excitation filter, a 420 nm suppression filter and UVFL 100/1.30 objective and an improved lens system to obtain strong fluorescence<sup>9,10</sup>. Under the microscope, various phages (T4, T7 and  $\phi 29$ ) and plasmids (pBR322 and pSC101) appeared as light-diffusion spots<sup>9,10</sup>. To take photomicrographs, the *Chlamydomonas* samples were fixed at various times after mixing, stained with DAPI dissolved in S buffer (20 mM Tris-HCl pH 7.6, 0.25 M sucrose, 1 mM EDTA, 0.6 mM spermidine, 7 mM  $\beta$ -mercaptoethanol and 1 mM phenylmethylsulphonyl fluoride) at  $5\ \mu\text{g}\ \text{ml}^{-1}$  for 10 min and squashed lightly or strongly on a glass slide. The preparations were fixed in 0.6% glutaraldehyde in S buffer for 30 s.

To observe the behaviour of cell nuclei, chloroplasts, chloroplast nucleoids and pyrenoids in gametes and zygotes, living and fixed cells were examined by phase-contrast and epifluorescent microscopy. The behaviour of the chloroplast during zygote formation (observed by phase-contrast microscopy) in

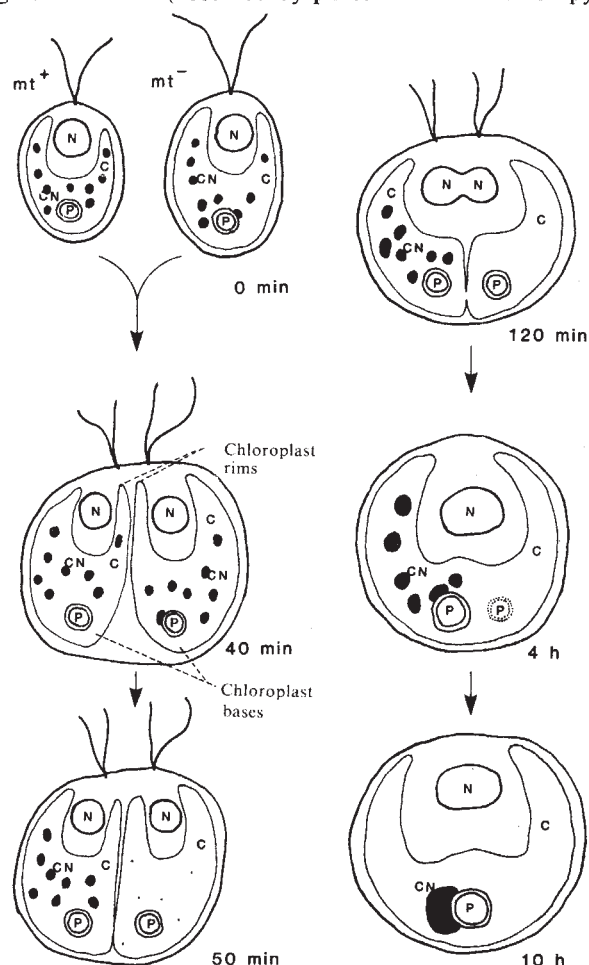
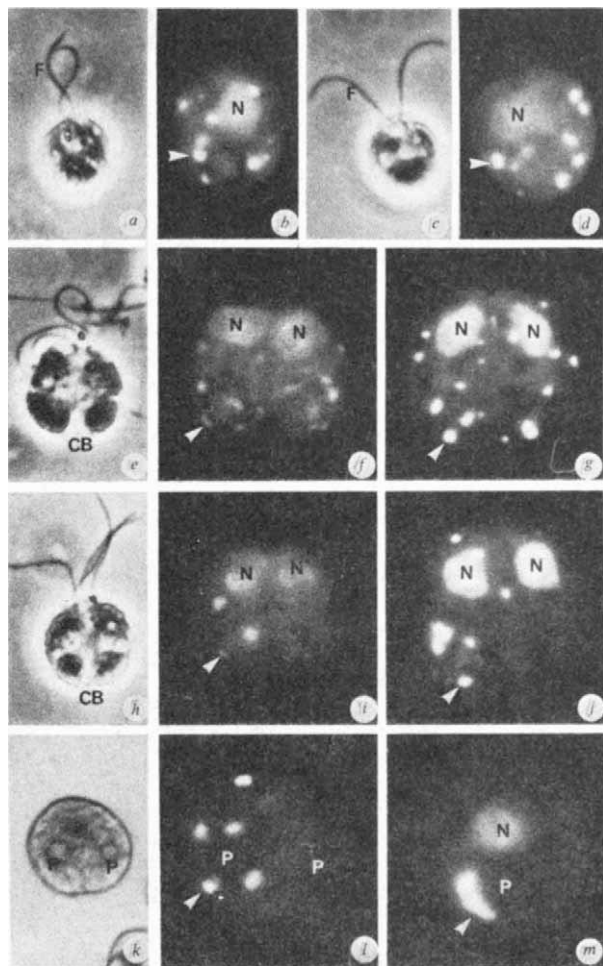


Fig. 1 Diagrammatic summary of the major events (approximate times indicated) during the first 10 h after cell fusion in *C. reinhardtii*, showing fusion of cell nuclei (N; 120 min), chloroplast (C) fusion (120 min), preferential loss of chloroplast nucleoids (CN) and the pyrenoid (P; 4 h) of male gamete origin, and fusion of chloroplast nucleoids of female origin around the remaining pyrenoid (10 h).





**Fig. 2** Phase-contrast (a, c, e, h), light field (k) and epifluorescent (b, d, f, g, i, j, l, m) photomicrographs showing flagella (F), pyrenoid (P) and chloroplast nucleoids (arrowheads) in  $mt^+$  gametes (a, b),  $mt^-$  gametes (c, d) and zygotes (e–m) of *C. reinhardtii* at 40 min (e–g), 50 min (h–j), 5 h (k, l) and 10 h (m) after mating. The cells were fixed in 0.6% glutaraldehyde for 30 s and stained with DAPI for 10 min, then lightly (a–f, h, i, k, l) or heavily (g, j, m) squashed on a glass slide. a, c, e, h and k are phase-contrast or light field images and b, d, f, i and l are epifluorescent images observed in the same field of view. At 50 min after mating, when the chloroplast bases (CB) were close together, the chloroplast nucleoids in the chloroplast from one of the parents had disappeared, while those in the chloroplast from the other parent remained (i, j). These chloroplast nucleoids were in the half containing the iodine-stained larger pyrenoid after chloroplast fusion (k, l), and they finally fused to form one large chloroplast nucleoid around the remaining pyrenoid (m). a, c, e, h and k,  $\times 2,378$ ; b, d, f, g, i, j, l and m,  $\times 3,850$ .

this study was basically similar to that reported by Cavalier-Smith<sup>11</sup>. Each wild-type gamete ( $mt^+$ ,  $mt^-$ ) has two flagella 8  $\mu\text{m}$  long (Fig. 2a, c) and contains a cell nucleus of 2  $\mu\text{m}$  diameter and a cup-shaped chloroplast (Fig. 1). After DAPI staining, each chloroplast emitting red fluorescence for both  $mt^+$  (Fig. 2b) and  $mt^-$  (Fig. 2d) gametes contained small spherical chloroplast nucleoids 0.3–0.5  $\mu\text{m}$  in diameter which emitted stronger blue-white fluorescence than the cell nuclei. The mitochondrial nucleoids were very small and, as there were only a few, they were only rarely observed once the mitochondria had been squashed out from the cell. DAPI-stained gametes were treated on a glass slide with DNase and RNase as described previously<sup>6</sup>. DNase treatment caused the fluorescence in the chloroplast nucleoids to disappear completely, but RNase produced no change in the fluorescent intensity. Thus, DAPI specifically stains DNA.

At 40 min after mating of gametes of the two wild types, the newly formed zygotes are spherical, have four flagella (Fig. 2e)

and contain two cell nuclei (Fig. 2f, g) and two discrete chloroplasts. The chloroplasts are cup-shaped and lie side by side at the base of the cell with the open ends of the cups facing the flagella (Fig. 1). In preparations of lightly squashed cells (Fig. 2f, i), the outlines of cell nuclei and chloroplast nucleoids are obscure compared with those of heavily squashed preparations (Fig. 2g, j). In the lightly squashed preparation, the two chloroplast bases in most zygotes seem to be separate (Fig. 2e). Each chloroplast contains about 8–10 dispersed chloroplast nucleoids (Fig. 2f, g). At 50 min after mating, when the chloroplast bases containing the pyrenoids appear to be touching (Figs 1, 2h), the chloroplast nucleoids in the cup-shaped chloroplast from one parent are observed to have disappeared completely (right half in Fig. 2i, j). These chloroplast nucleoids either disappeared simultaneously, or first those at the cup-shaped chloroplast rims disappeared, followed by the nucleoids at the chloroplast bases. They remained unswollen and disappeared as though they had dissolved from their periphery towards the centre.

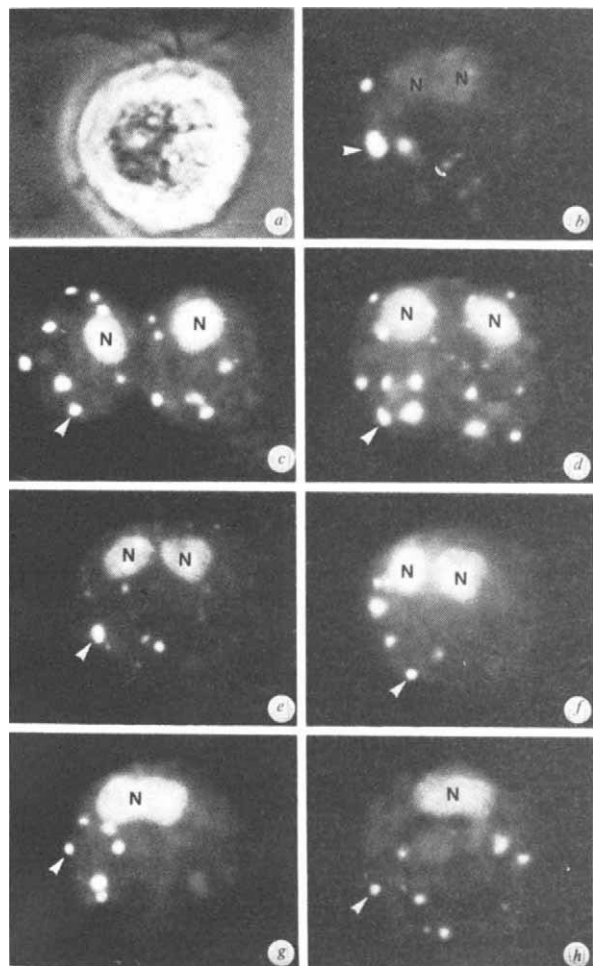
To determine whether or not the chloroplast nucleoids which disappeared were of  $mt^-$  origin, two crosses, that is,  $mt^+$  (gametes with normal flagella, wild type)  $\times$   $mt^-$  (gametes with short flagella, mutant) and  $mt^+$  (small gametes, wild type)  $\times$   $mt^-$  (large gametes, wild type) were prepared. In young zygotes of the first cross, the chloroplast nucleoids on the side with the short flagella disappeared while those on the side with the long flagella persisted (Fig. 3a, b). In young zygotes of the second cross, the chloroplast nucleoids on the 'small' side remained, while those on the 'large' side abruptly became small spherical light-diffusion spots (Fig. 3c–e) and disappeared completely before cell nuclear fusion (Fig. 3f), as observed for the normal cross (Fig. 2i, j). These results suggest that regardless of differences in cell size or flagellum length, chloroplast DNA from the  $mt^+$  parent is retained in the zygote while that from the  $mt^-$  parent is removed.

At 1–2 h after mating, the inner rims of the two chloroplasts which had previously been side by side, disappeared and cell nuclear fusion occurred (Figs 1, 3g). Next, the two chloroplasts fused together to form one large cup with two discrete pyrenoids (Fig. 1). The chloroplast nucleoids from the  $mt^+$  parent remained in one half of the chloroplast. At 4–5 h, the pyrenoid from the  $mt^-$  parent became reduced in size and disappeared, while the pyrenoid from the  $mt^+$  parent remained (Fig. 2k). Chloroplast nucleoids previously dispersed in the side containing the large pyrenoid (Fig. 2l), began to move towards the pyrenoid and fused together to form one large chloroplast nucleoid around it. By 12 h after mating, the large chloroplast nucleoid contained only one pyrenoid at its base and one to three chloroplast nucleoids were observed around this pyrenoid (Fig. 2m). These events have been summarized in Fig. 1.

The preferential destruction of the chloroplast nucleoids from the  $mt^-$  parent and the subsequent fusion of chloroplast nucleoids occur in most zygotes. However, in a few zygotes, the chloroplast nucleoids from the  $mt^-$  parent remained in the zygote even after fusion of the cell nuclei (Fig. 3h) and then fusion of the chloroplast nucleoids from both parents was observed.

The present epifluorescent microscopic observations are consistent with previous biochemical data which show that most of the chloroplast DNA from the  $mt^-$  parent is broken down by 6 h after mixing of the gametes while that from the  $mt^+$  parent persists<sup>3–5</sup>, except for the marked difference in the timing of preferential destruction of chloroplast DNA. The timing difference appears to be due to the degree of synchronization in zygote formation. As fusion of the gametes did not occur synchronously with their mixing, the resulting cell population contained male gametes, female gametes, and zygotes at various stages. Therefore, in the biochemical experiments, the preferential destruction of chloroplast DNA from male gametes should be markedly delayed compared with cytological observations.

Sager<sup>4</sup> suggested that a biochemical mechanism of maternal inheritance could be interpreted as a coupled methylation–restriction system: chloroplast DNA of female origin is heavily



**Fig. 3** Phase-contrast (a) and fluorescent (b–h) photomicrographs showing cell nuclei (N) and chloroplast nucleoids (arrowheads) in zygotes after the crosses  $mt^+$  (wild type)  $\times$   $mt^-$  (mutant with short flagella) (a, b) and  $mt^+$  (small gamete)  $\times$   $mt^-$  (large gamete) (c–g). The zygotes were fixed 2 (c), 30 (d), 45 (e), 50 (a, b, f) and 120 (g, h) min after mating and squashed lightly (a, b) or strongly (c–h) on the glass slide. The zygotes were stained with DAPI as described in Fig. 2 legend. a and b are the same field of view. In the cross  $mt^+$  (wild type)  $\times$   $mt^-$  (mutant), the chloroplast nucleoids on the side with the two long flagella persisted while chloroplast nucleoids on the other side (with two short flagella) began to disappear (a, b). In the cross  $mt^+$  (small gamete)  $\times$   $mt^-$  (large gamete), the female gametes were slightly smaller than the male gametes. The left half of the zygote is the area of  $mt^+$  origin, while the right half is of  $mt^-$  origin. Each chloroplast nucleoid of  $mt^-$  origin became smaller (e) and disappeared before fusion of the two cell nuclei (f), while the chloroplast nucleoids of  $mt^+$  origin persisted after cell nuclear fusion (g). After cell nuclear fusion, chloroplast nucleoids of  $mt^-$  origin were only rarely observed in the zygote (h). a,  $\times 3,760$ ; b–h,  $\times 4,125$ .

methylated during the first 6 h after mating and thus is not digested by the restriction enzyme, whereas that of male origin is unmethylated and thus is digested. Cavalier-Smith<sup>11</sup> found by electron microscopy that two chloroplasts in a zygote fused after the disappearance of the inner parts of the two cup-shaped chloroplast rims. As the present studies have shown that the chloroplast nucleoids of  $mt^-$  origin disappear while both chloroplast rims are still intact, the fusion of the cell nuclei and of the chloroplasts may not induce the activation of the restriction enzyme. The rapid destruction of the chloroplast nucleoids from the  $mt^-$  parent can account for the inheritance of maternal chloroplast DNA.

Rarely do chloroplast nucleoids from the  $mt^-$  parent persist after fusion of the two cell nuclei and of the chloroplasts (Fig. 3h)—this may explain the exceptional zygotes showing mater-

nal inheritance<sup>12</sup>. Within 10 h after fusion of the two gametes, the chloroplast nucleoids fuse to form one chloroplast nucleoid around the pyrenoid (Fig. 2m); the significance of the latter fusion remains unknown. Epifluorescent microscopy has shown that many mitochondrial nucleoids in yeast cells fuse to form one large nucleoid during the meiotic prophase<sup>13</sup>, when the recombination of yeast mitochondrial genes occurs with high frequency. By analogy, the fusion of chloroplast nucleoids in *C. reinhardtii* may be related to the recombination of chloroplast genes<sup>12</sup>. Further studies are in progress to elucidate the fate of large chloroplast nucleoids formed in the zygote by fusion of smaller nucleoids.

Received 2 April; accepted 20 May 1982.

1. Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **40**, 356–363 (1954).
2. Gillham, N. W. *Am. Nat.* **103**, 355 (1969).
3. Sager, R. & Lane, D. *Brookhaven Symp. Biol.* **23**, 495–505 (1972).
4. Sager, R. & Lane, D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2410–2413 (1972).
5. Sager, R. *Ann. N.Y. Acad. Sci.* **361**, 209–218 (1981).
6. Kuroiwa, T., Suzuki, T., Ogawa, K. & Kawano, S. *Pl. Cell Physiol., Kyoto* **22**, 322–338 (1981).
7. Chiang, K. S. *et al. Ann. N. Y. Acad. Sci.* **361**, 219–247 (1981).
8. Kuroiwa, T. & Suzuki, T. *Cell Struct. Function* **5**, 195–197 (1980).
9. Kuroiwa, T., Nishibayashi, S., Kawano, S. & Suzuki, T. *Experientia* **37**, 969–970 (1981).
10. Kuroiwa, T., Nishibayashi, S. & Kawano, S. *Proc. 13th Int. Bot. Congr. Sydney* **13**, 250 (1981).
11. Cavalier-Smith, T. *Nature* **228**, 333–335 (1970).
12. Adams, G. M. W., Vanwinkle-Swift, K. P., Gillham, W. W. & Boynton, J. E. *The Genetics of Algae* (ed. Lewin, R. A.) 69 (Blackwell Scientific, London, 1976).
13. Sando, N., Miyakawa, I., Nishibayashi, S. & Kuroiwa, T. *J. gen. appl. Microbiol.* **27**, 511–516 (1981).

## Stimulation of photosynthetic electron transport in a salt-tolerant plant by high chloride concentrations

Christa Critchley\*

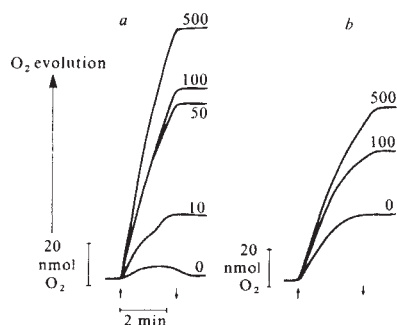
Department of Environmental Biology, Research School of Biological Sciences, Australian National University, Canberra, Australia

Halophytes are plants which tolerate high concentrations of electrolytes of which NaCl is normally the dominating salt. Compartmentation of the electrolytes within the cytoplasm has not been unequivocally resolved<sup>1</sup>. Several soluble enzymes from higher plant halophytes showed similar sensitivity to salinity when tested *in vitro* as did those from glycophytes<sup>2</sup>. Also, isolated mitochondria from halophytes and glycophytes showed similar inhibition of oxidative phosphorylation by NaCl *in vitro*<sup>3</sup>. On the other hand, chloride (plus the accompanying cations) has been shown to be accumulated in the chloroplasts of *Limonium vulgare*<sup>4–6</sup>, and it is also an essential cofactor in photosynthetic  $O_2$  evolution<sup>7–10</sup>. In the only report available on the effects of salt on chloroplasts from a halophyte<sup>11</sup>, it was demonstrated that differences exist in the response to salts of thylakoids from halophytic or mesophytic plants. I show here that thylakoids from mature leaves of the grey mangrove, *Avicennia marina*, require high chloride levels for photosynthetic electron transport around photosystem II and discuss the possibility of chloride accumulation into the chloroplast as an adaptational mechanism for salt tolerance.

Figure 1 shows the effect of added NaCl on  $O_2$  evolution in thylakoids isolated from leaves of *A. marina* with ferricyanide as electron acceptor. When these thylakoids were isolated and kept in the presence of 10 mM NaCl, 5 mM  $MgCl_2$ , virtually no  $O_2$  evolution was detectable without added NaCl in the assay. Rates of  $O_2$  evolution were improved and stabilized significantly by increasing the NaCl concentration (Fig. 1a). Preparation and storage of thylakoids in the presence of

\* Present address: Botany Department, The Faculties, Australian National University, PO Box 4, Canberra A.C.T. 2601, Australia.

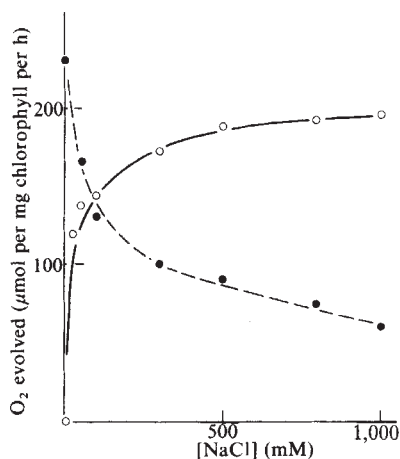




**Fig. 1** Kinetics of  $\text{O}_2$  evolution with  $1.5 \text{ mM } [\text{Fe}(\text{CN})_6]^{3+}$  as electron acceptor in thylakoids from *A. marina* isolated in the presence of: a,  $10 \text{ mM NaCl}$ ,  $5 \text{ mM MgCl}_2$ ; b,  $500 \text{ mM NaCl}$ ,  $5 \text{ mM MgCl}_2$ . Numbers on traces indicate concentrations in  $\text{mM}$  of added  $\text{NaCl}$  in the assay mix ( $50 \text{ mM Tricine-NaOH}$ ,  $\text{pH } 7.8$ ,  $1 \text{ mM MgCl}_2$ ,  $100 \text{ mM sucrose}$ )  $\uparrow$  light on;  $\downarrow$  light off. *A. marina* plants were grown in sand with  $50\%$  seawater (corresponding to  $\sim 250 \text{ mM NaCl}$  salinity) and supplied weekly with Hoagland's solution. They were kept in a growth cabinet with a  $25^\circ\text{C}/20^\circ\text{C}$  day/night temperature regime and a 12-h light period. Relative humidity was  $75\%$  and quantum density at leaf level  $\sim 400 \mu\text{E m}^{-2} \text{ s}^{-1}$ . Thylakoids were prepared by cutting four to six leaves ( $2\text{--}3 \text{ g}$  fresh weight) into  $2\text{-mm}^2$  pieces and, using mortar and pestle, grinding those with sand in a small quantity of grinding buffer ( $50 \text{ mM HEPES}$ ,  $\text{pH } 7.6$ ,  $5 \text{ mM MgCl}_2$ ,  $\text{NaCl}$ ,  $20\%$  polyethyleneglycol (PEG) 6000,  $5 \text{ mM dithiothreitol}$  (DTT). The homogenate was filtered through four layers of Miracloth and pelleted at  $200\text{g}$  for  $5 \text{ min}$ . The supernatant was centrifuged at  $1,000\text{g}$  for  $10 \text{ min}$ , the pellet resuspended in grinding buffer minus PEG and DTT and the suspension pelleted again for  $12 \text{ min}$  at  $1,500\text{g}$ . The thylakoid pellet was resuspended in grinding buffer minus PEG and DTT.

$500 \text{ mM NaCl}$  resulted in appreciable but unstable rates of  $\text{O}_2$  evolution when assayed without added  $\text{NaCl}$ . These rates were also significantly improved and stabilized by increased  $\text{NaCl}$  concentrations in the assay medium (Fig. 1b). The optimal salinity range for  $\text{O}_2$  evolution of isolated thylakoids *in vitro* was  $250\text{--}500 \text{ mM NaCl}$ . As shown in Fig. 2,  $\text{O}_2$  evolution was maintained at the same maximum level even up to  $1 \text{ M NaCl}$ . In contrast, these same high salt concentrations had an inhibitory effect on thylakoids isolated from cucumber leaves (Fig. 2), confirming recent results obtained by Baker with spinach thylakoids<sup>12</sup>.

Electron transport associated with photosystem I (PSI) in mangrove thylakoids did not exhibit the same dependence on



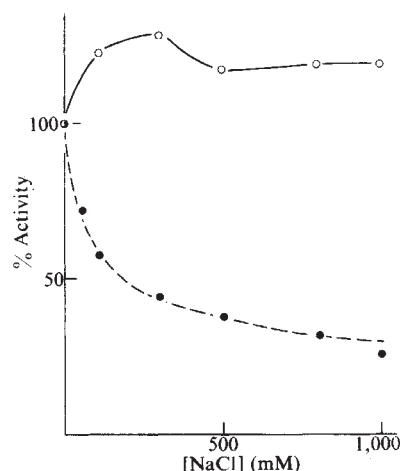
**Fig. 2** Effect of  $\text{NaCl}$  on  $\text{O}_2$  evolution with  $1.5 \text{ mM } [\text{Fe}(\text{CN})_6]^{3+}$  as electron acceptor in thylakoids from *A. marina* (○) and *C. sativus* (●). Electron transport activities were measured polarographically in a Clark-type  $\text{O}_2$  electrode in saturating white light provided by a slide projector. Assay temperature was  $25^\circ\text{C}$  and  $20\text{--}25 \mu\text{g ml}^{-1}$  chlorophyll was used in a total volume of  $3 \text{ ml}$ . The basic assay medium consisted of  $50 \text{ mM Tricine-NaOH}$ ,  $\text{pH } 7.8$ ,  $1 \text{ mM MgCl}_2$  and  $100 \text{ mM sucrose}$ .

**Table 1** Specificity of chloride effect

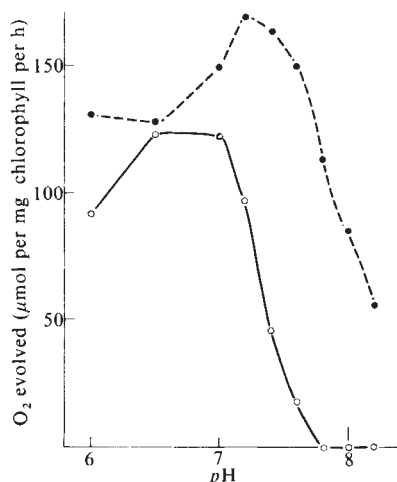
| Addition to assay                             | $\text{H}_2\text{O} \rightarrow \text{FeCN}$ (relative rates) |
|-----------------------------------------------|---------------------------------------------------------------|
| $500 \text{ mM NaCl}$                         | 100                                                           |
| $100 \text{ mM KCl}$                          | 78                                                            |
| $250 \text{ mM choline chloride}$             | 85                                                            |
| $100 \text{ mM Na}_2\text{SO}_4$              | 0                                                             |
| $500 \text{ mM KI}$                           | 0                                                             |
| $500 \text{ mM KBr}$                          | 0                                                             |
| $1 \text{ M sucrose}$                         | 15                                                            |
| $0.5 \text{ M sucrose} + 0.25 \text{ M NaCl}$ | 82                                                            |

high  $\text{Cl}^-$ . It did, however, show a similar resistance to very high salinities (Fig. 3). All above results are qualitatively very similar to those obtained by Hind *et al.*<sup>10</sup> with  $\text{Cl}^-$ -depleted spinach thylakoids. However, there is a concentration difference of two orders of magnitude in the effectiveness of  $\text{Cl}^-$ : whereas  $500 \text{ mM NaCl}$  was required for maximal steady  $\text{O}_2$  evolution rates in mangrove thylakoids,  $5 \text{ mM NaCl}$  was sufficient in spinach thylakoids. It was not possible to obtain whole chain electron transport ( $\text{H}_2\text{O} \rightarrow \text{MV}$ ) in thylakoids from *A. marina* in the conditions used. The electron transport chain between the two photosystems may be interrupted, presumably through inhibition or loss of one or more intermediate electron carriers. Ferricyanide reduction was very sensitive to (3-3',4'-dichlorophenyl)-1,1-dimethylurea (DCMU) but not to 2,5-dibromo-3-methyl-6-isopropyl-p-benzoquinone (DBMIB). Therefore, it appears that ferricyanide accepts electrons in these mangrove thylakoids between the primary acceptor of PSII, Q, and plastoquinone. This has also been found to occur in sonicated chloroplasts, where PSII directly reduces ferricyanide in a DBMIB-insensitive fashion<sup>13</sup>. Isolated thylakoids from *Euglena* showed somewhat similar properties<sup>14</sup>, including a positive response to chloride. Kelley and Izawa<sup>15</sup> found that their chloride-depleted spinach thylakoids were only inhibited 50% by DBMIB.

Various workers<sup>8,10,14</sup> observed a marked  $\text{pH}$  dependence of the  $\text{Cl}^-$  effect in their non-halophytic systems. A similar but much more dramatic effect of  $\text{pH}$  was found in thylakoids from *A. marina* in the absence of added  $\text{NaCl}$  (see Fig. 4). In the absence of added  $\text{NaCl}$ , appreciable rates of  $\text{O}_2$  evolution were observed only up to  $\text{pH } 7.2$ . The rates measured at  $\text{pH } 7.4$  and  $7.6$  were very unstable, and at  $\text{pH } 7.8$   $\text{O}_2$  evolution was completely inhibited. In the presence of  $100 \text{ mM NaCl}$ , rates of  $\text{O}_2$  evolution were generally higher, the optimum being around  $7.2\text{--}7.4$ .



**Fig. 3** Effect of  $\text{NaCl}$  on PSI activity in thylakoids from *A. marina* (○) and *C. sativus* (●) in  $50 \text{ mM Tricine-NaOH}$ ,  $\text{pH } 7.8$ ,  $1 \text{ mM MgCl}_2$ ,  $100 \text{ mM sucrose}$ ,  $0.1 \text{ mM 2,6-dichlorophenolindophenol}$  (DCIP),  $1 \text{ mM ascorbate}$ ,  $0.1 \text{ mM methylviologen}$ . Control activities, that is, 100% values, in  $\mu\text{mol per mg chlorophyll per h}$ : *A. marina* 150; *C. sativus* 386.



**Fig. 4** pH dependence of the  $\text{Cl}^-$  effect on  $\text{O}_2$  evolution in *A. marina* thylakoids. ○, No NaCl added; ●, 100 mM NaCl added. pH 6.0–6.5, 50 mM MES; pH 7.0–8.2, 50 mM Tricine. Other assay conditions as for Fig. 2.

Hind *et al.*<sup>10</sup> reported the effectiveness of  $\text{Cl}^-$  in preventing photoinhibition in  $\text{Cl}^-$ -deficient thylakoids. Similar experiments with mangrove thylakoids qualitatively confirm those earlier results. Pre-illumination with strong white light ( $1,800 \mu\text{E m}^{-2} \text{s}^{-1}$ ) of mangrove thylakoids in the absence of added NaCl resulted in decreasing rates of  $\text{O}_2$  evolution obtained on re-addition of  $\text{Cl}^-$  and ferricyanide such that after 4 min only 30% of the control activity remained. In contrast, pre-illumination in the presence of 100 mM NaCl largely prevented this loss of activity (data not shown).

Without  $\text{Cl}^-$  as the accompanying anion, none of the cations listed in Table 1 were effective in stimulating  $\text{O}_2$  evolution at the levels tested here; neither could iodide or bromide substitute for chloride. Osmotic effects were also not responsible because 1 M sucrose had hardly any influence on activity (Table 1). With choline as a nonpermeant cation,  $\text{Cl}^-$  was still fully effective.

A  $\text{Cl}^-$  requirement for PSII electron transport from water could only be observed in spinach or pea thylakoids after depletion by washings or ageing<sup>10,15,16</sup>. A concentration of 5–10 mM NaCl was sufficient to restore lost  $\text{O}_2$  evolution activity. The mangrove thylakoids used in the present study were not rendered  $\text{Cl}^-$  free by any of the above treatments and thus could not be considered  $\text{Cl}^-$  deficient by the same criteria. If chloroplasts of *A. marina* contain high  $\text{Cl}^-$  concentrations *in vivo*, they may therefore require these high  $\text{Cl}^-$  concentrations for  $\text{O}_2$  evolution *in vitro*. In that case, preparation and storage in 10 or 50 mM NaCl would render the thylakoids  $\text{Cl}^-$  deficient, this deficiency being further aggravated by actually diluting them into a  $\text{Cl}^-$ -free assay medium. The latter effect would then also be observed in thylakoids which were prepared and kept in 500 mM NaCl.

Chloroplasts from *Limonium vulgare* and *Tolypella intricata*, isolated by a non-aqueous technique, have indeed been shown to contain  $\text{Cl}^-$  concentrations<sup>4,5</sup> which were 10 times higher than that of comparable nonhalophytic species<sup>4</sup>. Using an entirely different approach and techniques, Ziegler and Lüttge<sup>6</sup> showed that  $^{36}\text{Cl}^-$ , when fed to intact plants of *Limonium vulgare*, was almost exclusively accumulated in the chloroplasts. It is therefore reasonable to suggest that an increased chloroplast  $\text{Cl}^-$  concentration has resulted in an increased  $\text{Cl}^-$  requirement for  $\text{O}_2$  evolution in at least some halophytes.

Most of the  $\text{Cl}^-$  necessary for  $\text{O}_2$  evolution in mangrove thylakoids is apparently not tightly bound to the membrane or one or more of its constituents, because low  $\text{Cl}^-$  concentrations during isolation and assay serve to inhibit  $\text{O}_2$  evolution without the need of any further depletion. The inhibition is also fully reversible and does not seem to result in any permanent damage of the membrane. It is therefore suggested that the effect is

based on ionic binding and dissociation of  $\text{Cl}^-$  to sites on the membrane and/or the  $\text{O}_2$  evolving enzyme. At high pH in the absence of sufficiently high  $\text{Cl}^-$  concentrations,  $\text{OH}^-$  ions could bind to these sites and prevent  $\text{O}_2$  evolution. Replacement by  $\text{Cl}^-$  ions would then restore activity.

The *in vitro* effect of chloride on photosynthetic electron transport in mangrove thylakoids could reflect a physiological requirement for high levels of chloride in this halophytic species, which in turn might be due to an adaptational process in which chloride is preferentially accumulated in chloroplasts. Whether the mechanism of  $\text{Cl}^-$  action is the same in both glycophytes and halophytes remains to be resolved, but the resolution is of fundamental interest to the explanation of concepts underlying adaptation processes. Further experiments should also reveal the usefulness of the mangrove thylakoids in the investigation of the still elusive mechanism of  $\text{O}_2$  evolution.

I thank Marilyn Ball, who suggested isolating thylakoids from mangrove leaves, for many stimulating discussions, Anne Gallagher for technical assistance, Barry Osmond for support and encouragement, and Ion Baianu, Govindjee and Martin Spalding for helpful discussions and critical reading of the manuscript. The manuscript was written while I was supported by NSF grant PCM 79-11148 to H. S. Gutowsky of the Department of Chemistry, University of Illinois, Urbana.

Received 6 October 1981; accepted 6 January 1982.

1. Flowers, T. J., Troke, P. F. & Yeo, A. R. A. *Rev. Pl. Physiol.* **28**, 89–121 (1977).
2. Greenway, H. & Osmond, C. B. *Pl. Physiol.* **49**, 256–259 (1972).
3. Flowers, T. J. *J. exp. Bot.* **101**, 101–110 (1974).
4. Larkum, A. W. D. *Nature* **218**, 447–449 (1968).
5. Larkum, A. W. D. & Hill, A. E. *Biochim. biophys. Acta* **203**, 133–138 (1970).
6. Ziegler, H. & Lüttge, U. *Planta* **74**, 1–17 (1967).
7. Warburg, O. & Lüttgens, W. *Naturwissenschaften* **32**, 301 (1944).
8. Gorham, P. R. & Clendenning, K. A. *Archs Biochem. Biophys.* **37**, 199–233 (1952).
9. Punnett, T. P. *Pl. Physiol.* **34**, 283–289 (1959).
10. Hind, G., Nakatani, H. Y. & Izawa, S. *Biochim. biophys. Acta* **172**, 277–289 (1969).
11. Wignarajah, K. & Baker, N. R. *Physiol. Pl.* **51**, 387–393 (1981).
12. Baker, N. R. *Pl. Physiol.* **62**, 889–893 (1978).
13. Böhme, H., Reimer, S. & Trebst, A. *Z. Naturforsch. b* **26**, 341–352 (1971).
14. Satoh, K., Katoh, S. & Takamiya, A. *Pl. Cell Physiol.* **11**, 453–466 (1970).
15. Kelley, P. M. & Izawa, S. *Biochim. biophys. Acta* **502**, 198–210 (1978).
16. Heath, R. L. & Hind, G. *Biochim. biophys. Acta* **172**, 290–299 (1969).

## Cloning of *Rhizobium meliloti* nodulation genes by direct complementation of $\text{Nod}^-$ mutants

Sharon R. Long\*, William J. Buikema & Frederick M. Ausubel

Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138, USA

Nitrogen-fixing root nodules are formed by the multi-step interaction of an endosymbiotic bacterium in the genus *Rhizobium* and a plant host of the legume family<sup>1</sup>. Identifying the bacterial and host genes involved in each step of this process is necessary to elucidate the molecular basis of the symbiosis, and is an important prerequisite for improving existing symbiotic associations or extending the range of symbioses. One way of identifying such genes is through the isolation and characterization of mutants that affect the symbiosis<sup>2–7</sup>. We have previously described the isolation of nodulation defective ( $\text{Nod}^-$ ) mutants of *Rhizobium meliloti*<sup>6,8</sup>. However, the functions of the nodulation (*nod*) genes defined by these mutations are not understood. As a first step in the molecular genetic characterization of *nod* genes, we describe here the cloning, by direct selection, of *nod*

\* To whom correspondence should be addressed at: Department of Biological Sciences, Stanford University, Stanford, California 94305, USA.

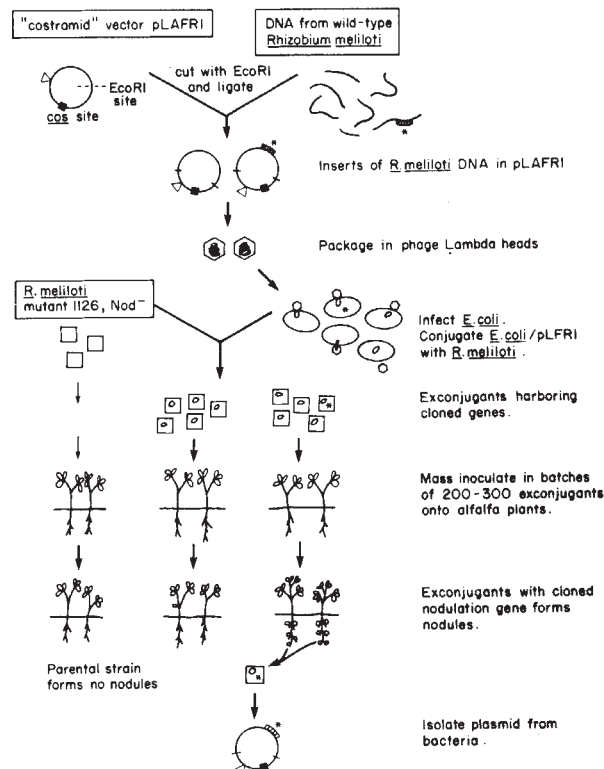


genes which complement the nodulation defect of a *Nod*<sup>-</sup> *R. meliloti* mutant. We have used a broad host range cosmid as a vector for large *R. meliloti* DNA inserts. This system should be generally useful for genetic studies of other Gram-negative bacteria which cannot easily be transformed; combined with the strategy of direct selection, it may be especially applicable to the study of virulence genes in pathogenic bacteria. Using the cloned *R. meliloti* nodulation gene obtained by this method, and a series of overlapping members of a cosmid clone bank extending from the *nif* locus, we have found that the gene(s) for nodulation are located within 30 kilobase pairs (kbp) of the *nif* loci, on the *R. meliloti* *nifK* side of the three structural genes for nitrogenase.

We have previously described<sup>6</sup> the isolation of Tn5-induced mutants of *R. meliloti* that did not form nodules on host alfalfa plants (*Nod*<sup>-</sup>), or that induced nodule formation but failed to fix nitrogen (*Fix*<sup>-</sup>). In the case of some Tn5-containing mutants, the mutated region of the genome was successfully cloned by genetic selection in *Escherichia coli* for the *R. meliloti* DNA fragment that contained the transposon (S.R.L., unpublished). However, in other mutants, this strategy was not possible because either (1) the Tn5 insertion was flanked by phage Mu DNA sequences<sup>6</sup>, resulting in difficulty in cloning the region, or (2) the mutation causing the *Nod*<sup>-</sup> or *Fix*<sup>-</sup> phenotype was found to be due to the insertion of an endogenous *R. meliloti* insertion sequence rather than a Tn5 insertion (ref. 9 and S.R.L., unpublished).

Among the symbiotic mutants with such complications were two independently isolated 'unreactive *Nod*<sup>-</sup>' mutants, 1027 and 1126, which did not curl root hairs of the host plant, an early stage of symbiotic interaction<sup>6,8</sup>. We therefore sought another method of cloning the genes for these mutations. It seemed possible that the host plant would be able to select rare *Nod*<sup>+</sup> *R. meliloti* cells from a large group of *Nod*<sup>-</sup> cells. In a reconstruction experiment, we found that a nodulation proficient strain of *R. meliloti* (1021) was able to induce nodules on plants even if outnumbered by 10<sup>4</sup> in a mixed inoculum with either mutant 1027 or 1126. Accordingly, we attempted to use plants to select directly for those *Nod*<sup>-</sup> cells that had received a cloned gene complementing the *Nod*<sup>-</sup> mutation. The outline of the selection scheme is shown in Fig. 1. Brewin *et al.*<sup>10</sup> used a similar strategy to obtain *Rhizobium leguminosarum* colonies which had received an indigenous conjugative *R. leguminosarum* plasmid and had thereby acquired a new nodulation range. Two considerations in carrying out such an experiment are that the cloning vector, containing wild-type *R. meliloti* DNA sequences, must be able to transfer into and replicate in *R. meliloti*, and that the clone bank should have large contiguous stretches of DNA inserted, to maximize the chance of finding intact functioning genetic units. We used a clone bank of *R. meliloti* DNA prepared in a wide host-range cosmid vector, pLAFR1<sup>11</sup>. The construction of pLAFR1 and its use in complementing auxotrophic mutations have been described elsewhere<sup>11</sup>.

The *R. meliloti*-pLAFR1 clone bank, derived from a partial *Eco*RI digest of nodulation-proficient strain 1021 DNA and maintained in *E. coli* HB101, was mass-conjugated into either *R. meliloti* 1027 or 1126, in a tri-parental mating using HB101/pRK2013 as a source of helper plasmid<sup>5</sup>. Tetracycline-resistant (*Tc*<sup>r</sup>) *R. meliloti* exconjugants were scraped together in batches of 200–300 colonies, suspended in sterile water, and each mixture was inoculated onto a group of five to eight aseptically grown alfalfa plants (*Medicago sativa* L. cv 'Iroquois'), each grown in a separate test tube on nitrogen-free nutrient agar<sup>6</sup>. Plants were observed after 4 weeks for the presence or absence of root nodules. If any one of the 200–300 colonies in an inoculation mix contained the complementing plasmid then all five to eight plants should have shown nodules; if no colony contained that cloned fragment, then none of the plants should have been nodulated by the mixture (excluding revertants of the *Nod*<sup>-</sup> mutants). Previous studies had shown that any particular auxotrophic genetic locus was represented



**Fig. 1** Selection scheme for obtaining cloned nodulation genes by complementation of *Nod*<sup>-</sup> *R. meliloti* mutants. The clone bank was constructed by partial digestion of wild-type *R. meliloti* 1021 DNA with *Eco*RI and insertion into the *Eco*RI site of pLAFR1<sup>11</sup>, a low copy number broad host range vector, which we call a 'cosmid', which contains a 1.6-kb *Bgl*II fragment, derived from pHC79<sup>14</sup>, containing  $\lambda$  *cos*, ligated into the unique *Bgl*II site in pRK290<sup>15</sup>. pLAFR1 is 21.6 kb, confers tetracycline resistance (*Tc*<sup>r</sup>), contains a unique *Eco*RI site, and can be mobilized from *E. coli* into *R. meliloti* by the specially constructed complementing plasmid pRK2013<sup>16</sup>. The *R. meliloti*-pLAFR1 ligated DNA mix was packaged in phage  $\lambda$  heads as described previously<sup>17</sup> using the  $\lambda$  lysogenic strains NS428 and NS433<sup>18</sup>. The mean insert size was 23.1 kb and the bank contained ~15,000 independent plasmids. The pLAFR1 clone bank, maintained in *E. coli* strain HB101<sup>19</sup>, was conjugated into *R. meliloti* strains 1027 and 1126 in a tri-parental mating as described elsewhere<sup>5</sup>. *Tc*<sup>r</sup> exconjugants were plated at a density of 200–300 per plate and 200–300 colonies were scraped and mixed together for sib-screens of nodulation ability on sterile alfalfa plants<sup>6</sup>. Nodules which formed on plants were removed, washed in 0.5% SDS and 10 mM NaCl, submerged for 5 min in 70% ethanol, washed twice in sterile H<sub>2</sub>O and squashed using a sterile glass rod into 1.0 ml of a solution containing 12% sucrose, 50 mM Tris, 5 mM EDTA, 10 mM NaCl, pH 7.5. Diluents of this mixture were plated on LB plates [10 g tryptone (Difco), 5 g yeast extract (Difco) 5 g NaCl, 5 g agar (Difco), pH 7.2], purified to single colonies, and tested for tetracycline resistance. Nodulated plants were assayed for nitrogenase activity using the acetylene reduction method as described previously<sup>5</sup>.

at a frequency of between 1 in 200 and 1 in 1,000 in the *R. meliloti*-pLAFR1 clone bank<sup>11</sup>, so we expected some mixtures of 200–300 plasmids to harbour the desired *nod* gene, and others to lack it. The results are shown in Table 1.

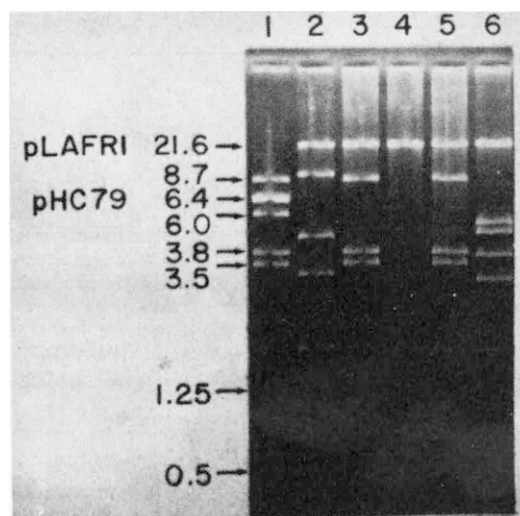
In conjugation A (Table 1), three inoculation mixtures produced no nodules on host plants, while inoculation with the other two batches of exconjugants resulted in nodule formation on roots of every plant. Essentially the same result was found with conjugation B, except for the formation of one 'rogue' nodule on a plant in group B-5, which was probably due to a *Nod*<sup>+</sup> revertant of strain 1027. Parental controls (plants inoculated with 1126 or 1027) were free of nodules.

To determine whether nodule formation by these groups of exconjugants was due to their containing a cloned gene that

complemented the *nod* lesion, we surface-sterilized nodules from infected plants, squashed them, isolated endosymbiotic bacteria, and analysed these for plasmid content. We found that any one nodule usually yielded a mixture of *R. meliloti* cells containing pLAFR1 clones and cells with no plasmid (as indicated by lack of the pLAFR1-coded tetracycline resistance). Also, in some nodules (about one-quarter of those analysed), two types of plasmid were found among all the colonies isolated from the nodule (Fig. 2, lanes 2–3 and lanes 5–6). One of the two showed no consistent pattern; however, the other was always pRmSL26 (Fig. 2, lanes 3, 5). This suggested that clone pRmSL26 contained a gene or genes used in nodulation, which restore to these 'unreactive' *Nod*<sup>−</sup> mutants the ability to curl root hairs and form nodules.

To test whether plasmid pRmSL26 was responsible for the restoration of the *Nod*<sup>+</sup> phenotype, pRmSL26 DNA was isolated from *R. meliloti* and used to transform *E. coli* strain HB101. HB101/pRmSL26 cells were then used to donate the clone into both mutant 1027 and 1126 recipients. Single exconjugant colonies were individually inoculated onto alfalfa plants and observed for nodule formation. In all cases, the parental *Nod*<sup>−</sup> strains without the conjugated clone were *Nod*<sup>−</sup>. Exconjugants that had received clone pRmSL26 were universally *Nod*<sup>+</sup>, and the plants nodulated by these strains had the same level of nitrogenase activity as plants nodulated with a wild-type strain.

It has recently been shown that nitrogenase genes in *R. meliloti* are located on a large indigenous 'megaplasmid' (refs 12, 13 and W.J.B. *et al.*, in preparation). Members of our laboratory have used cosmid cloning vectors to clone large overlapping regions of the megaplasmid extending approximately 50 Kb on each side of the nitrogenase (*nif*) genes. Figure 3 shows the location and map of two of these clones extending on the *nifK* side of the *nif* loci. When DNA from two such clones, pRmBE2 and pRmWB726, was digested with *Eco*RI and compared by electrophoresis with pRmSL26, it appeared that there were several restriction fragments in common between them (Fig. 2). From the functional complementation tests described above, we knew that pRmSL26 contained nodulation gene(s) identified by mutants 1027 and 1126. This could not be done directly with clone pRmWB726 because it was constructed using a cosmid (pHC79) based on a ColE1 replicon which will not replicate in *Rhizobium* cells. To determine whether the sequences in pRmWB726 correspond to genes which had been mutated in mutants 1027 and 1126, we used pRmWB726 DNA as a hybridization probe to DNA from these two strains. In both 1027 and 1126 it was found that the 8.7-kb



**Fig. 2** *Eco*RI digests of plasmids isolated from nodules (lanes 2–6) compared with an extended cloned sequence from the *R. meliloti* megaplasmid contained on the cosmid clone pRmWB726. Plasmids were isolated from *R. meliloti* (pLAFR1) or *E. coli* (pRmWB726) using an SDS-NaOH mini-preparation technique<sup>20</sup>. Lane 1, *Eco*RI digest of pRmWB726, a cosmid carrying a 23.7-kb insert of *R. meliloti* DNA in pHC79<sup>14</sup>. The insert in pRmWB726 begins ~21 kb from the nitrogenase region on the *R. meliloti* megaplasmid (W.J.B., unpublished; see Fig. 3). Lanes 2–4, *Eco*RI digests of three pLAFR1 cosmid plasmids isolated from a nodule on a plant in group A-4 (Table 1; complementation of strain 1126). The plasmid in lane 4 appears to contain the cosmid vector alone, apparently a doublet which would be of an appropriate size to be packaged. Other colonies from this nodule contained the same plasmid seen in lane 3. Lanes 5, 6, *Eco*RI digests of two plasmids isolated from a single nodule on a plant in group B-2. The plasmid in lanes 3 and 5 appeared to be identical; one was chosen and designated pRmSL26.

**Table 1** Complementation of *Nod*<sup>−</sup> *R. mutants* using a cosmid clone bank

| Conjugation A            |                            |                                             |
|--------------------------|----------------------------|---------------------------------------------|
| Donor: HB101/pLAFR1 bank |                            | Recipient: 1126 ( <i>Nod</i> <sup>−</sup> ) |
| Exconjugant mixture      | No. of exconjugants tested | Nodulated plants/total plants               |
| A-1                      | 200                        | 0/6                                         |
| A-2                      | 200                        | 0/7                                         |
| A-3                      | 200                        | 5/5                                         |
| A-4                      | 200                        | 7/7                                         |
| A-5                      | 200                        | 0/7                                         |
| Conjugation B            |                            |                                             |
| Donor: HB101/pLAFR1 bank |                            | Recipient: 1027 ( <i>Nod</i> <sup>−</sup> ) |
| Exconjugant mixture      | No. of exconjugants tested | Nodulated plants/total plants               |
| B-1                      | 300                        | 0/8                                         |
| B-2                      | 300                        | 7/7                                         |
| B-3                      | 300                        | 8/8                                         |
| B-4                      | 300                        | 0/8                                         |
| B-5                      | 300                        | 1/8*                                        |

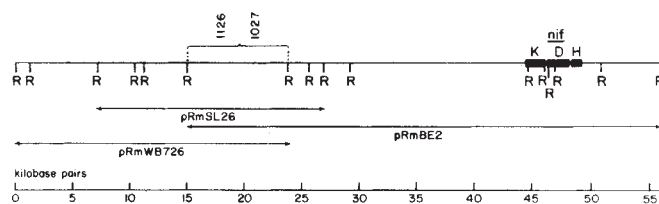
\* The nodulated plant contained a single nodule.

fragment present in both pRmSL26 and pRmWB726 was altered in size due to an insertion. Therefore, through functional and physical tests, we have assigned these nodulation functions to the region of the 8.7-kb *Eco*RI fragment on the megaplasmid, as shown in Fig. 3. As the insertions could result in polar effects on the adjacent genes, we cannot determine whether the nodulation genes are actually in the 8.7-kb fragment or an adjacent fragment. Other laboratories have also recently obtained evidence of the close linkage of *nod* and *nif* genes in *Rhizobium* species (refs 12, 13 and A. Johnston, personal communication).

We do not know why bacteria lacking pLAFR1 plasmids were common inside the selected nodules. However, we have observed that, while RP4 or its related plasmids are very stable in free-living cultures of *R. meliloti*, they are often lost if the *Rhizobium* which harbour them are inoculated onto plants and re-isolated from nodules. It is possible that because the *nod* genes encoded on pRmSL26 are used at an early stage of the symbiosis, they are not selected for during the proliferation of bacteria which follows successful infection, and are consequently lost. The possibility that such strains do not select for maintenance of the plasmid because the wild-type insert on pLAFR1 has recombined into the chromosome replacing the mutated *nod* gene is unlikely, because in the case of strain 1126, which contains Tn5 plus Mu sequences inserted in the 8.7-kb *Eco*RI fragments, the Tn5-encoded drug resistance for kanamycin/neomycin was present in all cells derived from the nodules. This test could not be carried out with strain 1027 because the mutation in the 8.7-kb fragment is not due to insertion of Tn5 but to insertion of a native *R. meliloti* insertion sequence, ISRM1, which bears no known marker.

We thank A. Friedman, C. Earl, G. Ruvkun, S. Brown for their help in this work, and R. Hyde for help in preparing the





**Fig. 3** Map of a region of the *R. meliloti* megaplasmid showing the nitrogenase genes (*nif*), *EcoRI* sites (R) and sequences contained in pRmSL26, pRmBE2 and in pRmWB726. The latter two plasmids were isolated from a cosmid clone bank in vector pHC79<sup>14</sup>. Clone pRmBE2 was identified by hybridization with pRmR2, which carries the *nifH* gene of *R. meliloti*. pRmWB726 was isolated on the basis of homology to the terminal 8.7-kb *EcoRI* fragment of pRmBE2.

manuscript. S.R.L. was supported by a postdoctoral fellowship from the National Institutes of Health; additional support came NSF grant PCM8104492 awarded to F.M.A.

Received 26 January; accepted 6 May 1982.

1. Vincent, J. M. in *The Biology of Nitrogen Fixation* (ed. Quispel, A.) 265–341 (North-Holland, Amsterdam, 1974).
2. Beringer, J. E., Beynon, J. L., Buchanan-Wollaston, V. & Johnston, A. W. B. *Nature* **276**, 633–634 (1978).
3. Johnston, A. W. B. *et al.* *Nature* **276**, 634–663 (1978).
4. Buchanan-Wollaston, V., Beringer, J. E., Brewin, N. J., Hirsch, P. R. & Johnston, A. W. B. *Molec. gen. Genet.* **178**, 185–190 (1980).
5. Ruvkun, G. B. & Ausubel, F. M. *Nature* **289**, 85–88 (1981).
6. Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E. & Ausubel, F. M. *J. Bact.* **149**, 114–122 (1982).
7. Rolfe, B. G. *et al.* in *Current Perspectives in Nitrogen Fixation* (eds Gibson, A. H. & Newton, W. E.) 142–145 (Australian Academy of Sciences, Canberra, 1981).
8. Long, S. R., Meade, H. M., Brown, S. E. & Ausubel, F. M. in *Genetic Engineering—Plant Sciences* (ed. Panopoulos, N. J.) 129–143 (Praeger, New York, 1981).
9. Ruvkun, G. B., Long, S. R., Meade, H. M. & Ausubel, F. M. *Cold Spring Harb. Symp. quant. Biol.* **45**, 492–499 (1981).
10. Brewin, N. J., Beringer, J. E. & Johnston, A. W. B. *J. gen. Microbiol.* **120**, 413–420 (1980).
11. Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J. & Ausubel, F. M. *Gene* (in the press).
12. Banfalvi, Z. *et al.* *Molec. gen. Genet.* **184**, 318–325 (1981).
13. Rosenberg, C., Boistard, P., Denarie, J. & Casse-Delbart, F. L. *Molec. gen. Genet.* **184**, 326–333 (1981).
14. Hohn, B. & Collins, J. *Gene* **11**, 291–298 (1980).
15. Ditta, B., Stanfield, S., Corbin, D. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7347–7351 (1980).
16. Figurski, D. H. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1648–1652 (1979).
17. Hohn, B. *Meth. Enzym.* **68**, 299–309 (1979).
18. Sternberg, N., Tiemeier, D. & Enquist, L. *Gene* **1**, 255–280 (1977).
19. Boyer, H. B. & Roulland-Dussoix, D. *J. molec. Biol.* **41**, 459–472 (1969).
20. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989–2998 (1981).

## Plasmids from *Staphylococcus aureus* replicate in yeast *Saccharomyces cerevisiae*

R. Goursot, A. Goze, B. Niaudet & S. D. Ehrlich

Institut de Recherche en Biologie Moléculaire, CNRS, Université Paris VII, 2, place Jussieu, 75251 Paris Cedex 05, France

It is known that some plasmids, such as RP4, can replicate in many Gram-negative bacteria<sup>1</sup>. Certain small *Staphylococcus aureus* plasmids have an even broader host range, being able to replicate in not only phylogenetically distant Gram-positive bacteria such as *Bacillus subtilis*<sup>2</sup> or *Streptococcus pneumoniae*<sup>3</sup>, but also in the Gram-negative bacterium *Escherichia coli*<sup>4</sup>. Here we have examined whether these plasmids can also replicate in a lower eukaryote, the yeast *Saccharomyces cerevisiae*. For this purpose we constructed hybrids between a *S. aureus* plasmid pC194 and an *E. coli* plasmid Yip5, which carries a *ura-3* gene easy to select for in yeast but cannot replicate in this host. We found that the hybrids transformed

yeast with high efficiency (as did hybrids between Yip5 and three other *S. aureus* plasmids); were maintained extrachromosomally in yeast; and were not modified during residence in yeast. We conclude from this evidence that *S. aureus* plasmids can replicate in yeast, which raises the questions of whether the replication signals used by prokaryotes and eukaryotes are similar, and how far up the phylogenetic tree the organisms still able to be hosts to *S. aureus* plasmids may be.

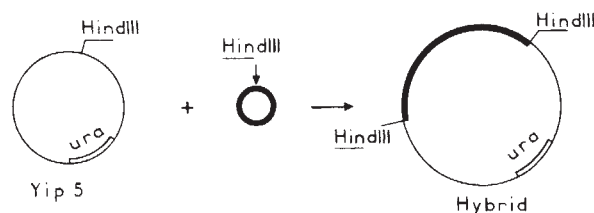
It has been observed that a yeast *ura-3* strain YNN27 cannot be transformed by the plasmid Yip5, which consists of pBR322 carrying a 1.1-kilobase (kb) yeast segment containing the *ura-3* gene. On the other hand, hybrids between Yip5 and replicons functional in yeast can transform YNN27 to uracil independence<sup>5</sup>. To test whether a plasmid from *S. aureus*, pC194<sup>6</sup>, can replicate in yeast we constructed hybrids between this plasmid and Yip5, as shown in Fig. 1. Two hybrids, pHV228 and pHV229, were of the expected structure, as judged by their size and restriction analysis. They differed in the relative orientation of pC194 and Yip5.

pHV228 and pHV229 were used to transform YNN27 cells to a *Ura*<sup>+</sup> phenotype. In a control experiment the cells were transformed with plasmid pG9, which is a hybrid containing the 1.1-kb *ura-3* segment linked to a part of the yeast 2  $\mu$  plasmid<sup>7</sup>. The results are shown in Table 1. All three plasmids transformed YNN27 with a frequency of well over 10<sup>3</sup> transformants per  $\mu$ g of DNA. No transformants were obtained with 1.4  $\mu$ g of the Yip5 plasmid.

Several lines of evidence indicate that pHV228 and pHV229 replicate autonomously in yeast. First, the *Ura*<sup>+</sup> phenotype of transformants is quite unstable; growth for five generations without selective pressure results in the appearance of 90% *Ura*<sup>−</sup> cells. This frequency is twice that observed with the transformants obtained with pG9 and much lower than that observed after insertion of the transforming plasmid into the chromosome (0.1–1% per cell division<sup>8</sup>).

Second, *Ura*<sup>+</sup> transformants were analysed by Southern-type hybridization. For this purpose, total DNA was extracted from yeast cells transformed with pHV228 or pHV229 and grown for some 50 generations in selective medium lacking uracil. It was analysed by electrophoresis after cleavage with enzymes that do not cut the plasmids, or with enzymes that cut the plasmids once. DNA was transferred to a nitrocellulose support and hybridized with the probe obtained by *in vitro* labelling of plasmid pHV33, a hybrid of pBR322 and pC194<sup>9</sup>. The results obtained with DNA extracted from cells containing pHV228 are shown in Fig. 2. An identical result was obtained when the cells contained pHV229.

A complex pattern, due to the existence of different molecular forms of the plasmid in the preparation, was observed with intact DNA. An identical pattern was seen with DNA cleaved with enzymes that do not cut pHV228. Hybridization at the position corresponding to that of linearized plasmid was observed when DNA was cleaved with enzymes that cut pHV228 once. Overexposure of the autoradiogram, which would have enabled us to detect a hybridization signal about 100 times smaller than that observed, failed to reveal any additional bands. In a control experiment no hybridization between pHV228 and the untransformed yeast DNA was observed.



**Fig. 1** Construction of hybrid plasmids. pC194, pC221, pC223 and pUB112<sup>11</sup> were inserted into Yip5 as shown in the scheme. pBR322, *ura-3* and *S. aureus* sequences are represented by thin, double and thick lines, respectively. All of the hybrid plasmids conferred ampicillin resistance and uracil independence on *E. coli* strain DB6656 *pyrF::Mu trp<sub>am</sub> lacZ<sub>am</sub> hsdR* (ref. 15).

**Table 1** Transformation of yeast YNN27 to uracil independence

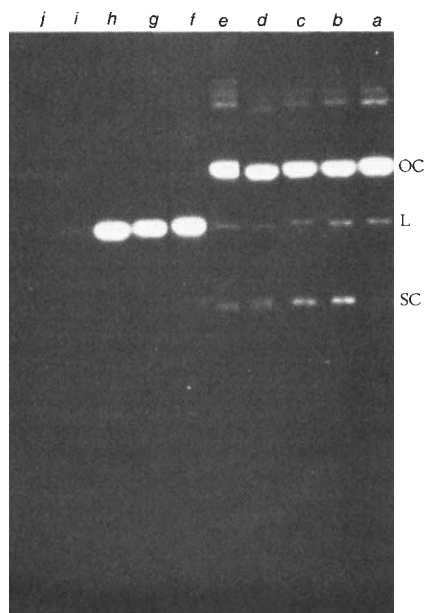
| Plasmid | Structure                    | Transforming efficiency<br>(no. of transformants<br>per $\mu\text{g}$ of DNA) |
|---------|------------------------------|-------------------------------------------------------------------------------|
| Yip5    | pBR322- <i>ura-3</i>         | <1                                                                            |
| pHV228  | pBR322- <i>ura-3</i> -pC194  | $3.5 \times 10^3$                                                             |
| pHV229  | pBR322- <i>ura-3</i> -pC194  | $7 \times 10^3$                                                               |
| pG9     | ColE1- <i>ura-3</i> -2 $\mu$ | $4 \times 10^3$                                                               |

Polyethylene glycol mediated transformation was performed as described elsewhere<sup>7</sup>.

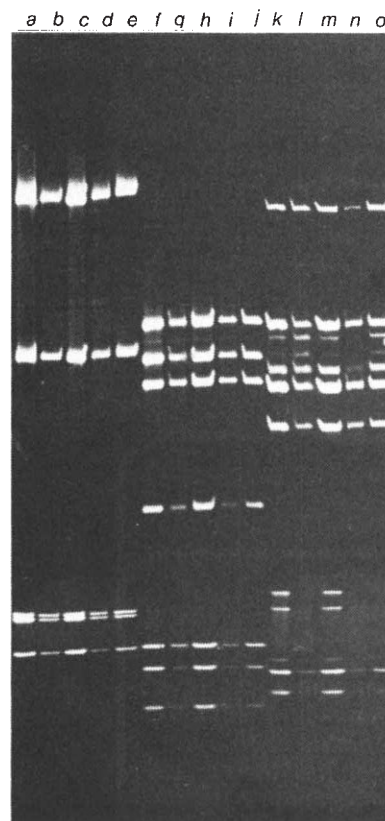
Had pHV228 and pHV229 been integrated into the chromosome, a single hybridization band rather than the observed pattern (Fig. 2) would have been obtained when the DNA extracted from Ura<sup>+</sup> transformants was cleaved with enzymes not cutting within the plasmids. Two hybridization bands instead of one would be obtained with enzymes which cut once within the plasmids. Our data therefore show that pHV228 and pHV229 are present in yeast as extrachromosomal elements.

The presence of autonomous plasmids in yeast cells transformed with pHV228 and pHV229 was also indicated by the fact that the DNA extracted from Ura<sup>+</sup> clones transformed competent *E. coli* cells to ampicillin resistance (Ap<sup>R</sup>), and *B. subtilis* to chloramphenicol resistance. This would not be expected if pHV228 and pHV229 were integrated into the yeast chromosome.

Two plasmids, pHV230 and pHV231, were prepared from representative Ap<sup>R</sup> clones, obtained by transforming *E. coli* with DNA extracted from yeast carrying plasmids pHV228 and pHV229, respectively. Analysis of pHV228, pHV229, pHV230 and pHV231 showed that the plasmids incurred no change during their residence in yeast: (1) the four plasmids were of the same size, had an identical restriction pattern when cleaved with *AluI*, *HaeIII*, *HinfI* and *TaqI*, and transformed yeast, *E.*



**Fig. 2** Hybridization analysis of DNA extracted from YNN27 cells transformed with pHV228. Total DNA was cleaved with enzymes *Bgl*II (b), *Hpa*I (c), *Sst*I (d) and *Sst*II (e) which do not cut pHV228 and with *Bam*HI (f), *Eco*RI (g) and *Sma*I (h) which cut it once. The samples, together with intact yeast (a) and pHV228 DNA (j) and pHV228 DNA cut with *Bam*HI (i), were separated by electrophoresis on a 0.5% agarose gel, transferred to a nitrocellulose support and hybridized with pHV33 DNA labelled by nick-translation. The three fastest migrating bands in lanes a-e correspond to supercoiled (SC), linear (L, arising presumably during DNA isolation) and nicked form of plasmid monomer (OC), the remaining bands to plasmid oligomers. The band in lanes f-i corresponds to linear plasmid monomer.



**Fig. 3** Comparison of authentic pC194 and four pC194-like plasmids by restriction analysis. The five DNAs were cleaved with *Taq*I (lanes a-e), *Mbo*II (lanes f-j) or *Alu*I (lanes k-o) and analysed on a polyacrylamide gel (top: 3%; bottom: 8%). The order of DNAs on the gel is authentic plasmid, plasmid excised from pHV228, pHV229, pHV230 and pHV231, respectively, for each restriction enzyme. Degradation with *Alu*I was incomplete.

*coli* and *B. subtilis* with the same efficiency; (2) when labelled by nick translation they hybridized with the same *Bgl*II and *Hpa*I segments of nontransformed YNN27 DNA as did the parental Yip5; (3) excision of pC194 from hybrid plasmids yielded four Yip5-like plasmids unable to transform yeast, and four pC194-like plasmids indistinguishable from the authentic pC194 by restriction analysis, which we estimate would have detected modifications greater than some 20–30 base pairs (see Fig. 3).

The copy number of hybrid plasmids was estimated by the spot hybridization technique<sup>10</sup>, using DNA extracted from yeast transformed with pHV228 or pHV229 and the labelled pHV33 probe. 0.5–2 plasmid copies per haploid genome were found using as a hybridization standard the mixture of DNA prepared from non-transformed YNN27 cells and known amounts of pHV228 or pHV229 DNA.

To establish whether *S. aureus* plasmids other than pC194 replicate in yeast we constructed hybrids between Yip5 and pC221, pC223 and pUB112<sup>11</sup> according to the scheme represented in Fig. 1. An additional *S. aureus* plasmid, pUB110<sup>12</sup>, devoid of *Hind*III sites but carrying a unique *Bam*HI site, was inserted into the *Bam*HI site of Yip5. The structure of the hybrid plasmids was as expected, as judged by their size and the digestion pattern with *Bam*HI or *Hind*III and at least one multi-cutting enzyme.

The constructed plasmids were used to transform YNN27 yeast cells. Over  $10^3$  transformants per  $\mu\text{g}$  of DNA were obtained in all cases, except that of the hybrid between Yip5 and pUB110, where a similar number of very small colonies, in addition to 1–10 larger colonies, per  $\mu\text{g}$  of DNA were obtained. This result was not due to the presence of a transformation inhibitor in the DNA preparation, because pG9 added



to the same DNA transformed efficiently. The very small colonies obtained with the YIp5-pUB110 hybrid are also regularly observed with YIp5 DNA; they did not grow on the selective medium on re-streaking, and are therefore likely to be abortive transformants. On the other hand, larger colonies obtained with this hybrid grew when streaked on the medium lacking uracil. The efficiency with which they were obtained corresponds to that observed for plasmids that cannot replicate in yeast, but can become inserted into the yeast chromosomes<sup>13</sup>. These results indicate that pC221, pC223 and pUB112 can replicate in yeast, whereas pUB110 cannot.

Plasmid pC194, originally isolated from *S. aureus*<sup>6</sup>, replicates in phylogenetically very distant prokaryotes<sup>2,4</sup>, which suggests that it can function in most prokaryotes and, as shown in the present work, in a lower eukaryote. This implies that either the prokaryote and lower eukaryote DNA replicating machineries have common features, enabling them to handle the same replicons, or the mode of pC194 replication is different in different prokaryote and eukaryote hosts. The study of the

relation between the structure and function of this replicon should enable us to distinguish between these alternatives.

Several other *S. aureus* plasmids appear to replicate in yeast, while one, pUB110, does not, as may be judged by the ability of a plasmid joined to YIp5 to transform yeast efficiently. The plasmids that replicate in yeast belong to different *S. aureus* incompatibility groups<sup>14</sup>, which suggests that their replication modes are different. Comparative analyses of these plasmids may reveal which replication signals enable them to function in phylogenetically very distant hosts. A question of some interest is whether organisms even more distant than those tested until now still recognize the same replication signals.

We thank R. Davis for strains, V. Ribes and M. Guérineau for initiating us in yeast transformation, and A. Dedieu and E. Pierre for constructing most of the hybrids used in this work. R.G. was the recipient of a fellowship from the DGRST; A.G. and S.D.E. are on the INSERM staff. This work was supported by grants ATP 72-79-104 to S.D.E. and CRL 1982 Com. 1 to A.G. from INSERM.

Received 29 March; accepted 2 June 1982.

1. Reaney, D. C. *Bact. Rev.* **40**, 552-590 (1976).
2. Ehrlich, S. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1080-1083 (1977).
3. Barany, F. & Alexander, T. J. *Bact.* **144**, 698-709 (1980).
4. Goze, A. & Ehrlich, S. D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7333-7337 (1980).
5. Stinchcomb, D. T., Thomas, M., Kelly, J., Selker, E. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4559-4563 (1980).
6. Iordanescu, S. J. *Bact.* **124**, 597-601 (1975).

7. Gerbaud, C. et al. *Gene* **5**, 233-253 (1979).
8. Hinnen, A. & Meyback, B. *Curr. Topics microbiol. Immun.* **96**, 101-117 (1981).
9. Primrose, S. B. & Ehrlich, S. D. *Plasmid* **6**, 193-201 (1981).
10. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
11. Novick, R. J. *Bact.* **127**, 1177-1187 (1976).
12. Gryczan, T. J., Contente, S. & Dubnau, D. J. *Bact.* **134**, 316-329 (1978).
13. Beggs, J. D. *Nature* **275**, 104-108 (1978).
14. Iordanescu, S., Surdeanu, M., Della Latta, P. & Novick, R. *Plasmid* **1**, 468-479 (1978).
15. Bach, M. L., Lacroute, F. & Botstein, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 386-390 (1979).

## Correlations between heterozygosity and evolutionary rate of proteins

D. O. F. Skibinski

Department of Genetics, University College of Swansea, Singleton Park, Swansea SA2 8PP, UK

R. D. Ward

Department of Human Sciences, University of Technology, Loughborough, Leicestershire LE11 3TU, UK

**Amino acid sequence analysis and electrophoretic analysis of protein variation have been widely used for measuring genetic differentiation between diverging lines of descent. The problem of identifying the causes of protein evolution and divergence is, however, largely unsolved. Using data extracted from a collection of allozyme surveys of vertebrate species, we report here that protein heterozygosity is correlated with both protein genetic distance and protein amino acid substitution rate. Proteins with high heterozygosity are shown to have evolved at a faster rate than proteins with low heterozygosity. Evolutionary rate estimates obtained from the two techniques of amino acid sequencing and electrophoretic analysis are highly correlated over those proteins assayed by both methods. The observed relationship between protein heterozygosity and genetic distance cannot be completely accounted for by neutral theory, but suggests that a substantial proportion of genetic distance accumulated between species is the result of selective substitution.**

Electrophoretic studies provide information on a variable that is of considerable importance for evolutionary theory. This is the genetic variation between individuals within a population or species. Such variation is widely assumed to be important as a prerequisite for evolutionary change; moreover, differences in genetic variability might explain or be used to predict differences in rates of evolution or divergence, higher variability

being associated with more rapid evolution<sup>1,2</sup>. Unfortunately, it is difficult to obtain direct comparisons between measures of protein variation such as observed or expected heterozygosity and amino acid substitution rates because sequence data are not readily available for the majority of enzymes or proteins most usually screened in individual electrophoretic surveys. Also, observed associations between protein variation and electrophoretic genetic distance estimates<sup>3,4</sup> are easily open to misinterpretation<sup>2,5</sup>, particularly when limited data are available for analysis.

In the present study, analyses are made on pooled data from a large number of allozyme surveys of fishes, amphibians, reptiles, birds and mammals. Most sources of data and the criteria used for their selection are given in ref. 5. The justification for pooling is that different vertebrate classes have in the past shown similar trends following various types of analysis<sup>6-9</sup>. In each survey, all possible pairwise interspecies comparisons were used to calculate the variables defined below. Let  $x_{ij}$  and  $y_{ij}$  be the frequencies of the  $i$ th allele at the  $j$ th locus in species  $X$  and  $Y$ , respectively. Then if  $a$  is the number of alleles,  $H_j$ , defined as

$$\left(2 - \sum_{i=1}^a x_{ij}^2 - \sum_{i=1}^a y_{ij}^2\right)/2,$$

is the arithmetic mean of the expected heterozygosity, according to Hardy-Weinberg, at the  $j$ th locus, and  $I_j$ , defined as

$$\sum_{i=1}^a x_{ij}y_{ij} / \left(\sum_{i=1}^a x_{ij}^2 \sum_{i=1}^a y_{ij}^2\right)^{1/2},$$

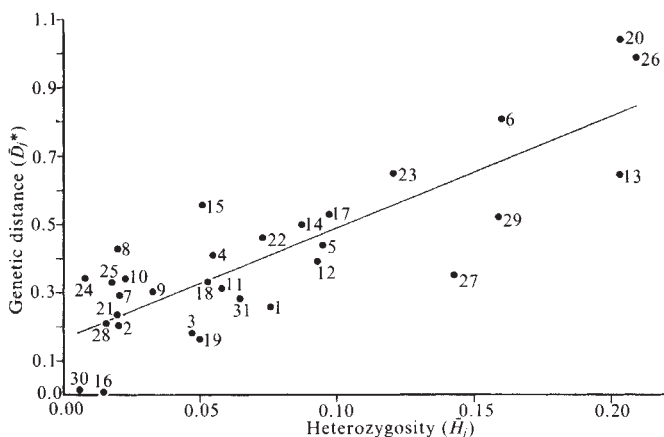
is the genetic identity value for  $X$  and  $Y$  at the  $j$ th locus. The normalized genetic identity between  $X$  and  $Y$  is defined as  $I = (J_{XY}) / (J_X J_Y)^{1/2}$ , where  $J_{XY}$ ,  $J_X$  and  $J_Y$  are arithmetic means of  $\sum x_{ij}y_{ij}$ ,  $\sum x_{ij}^2$  and  $\sum y_{ij}^2$  over all loci<sup>10</sup>.  $D$ , defined as  $-\log_e I$ , can be regarded as an estimate of the accumulated number of gene substitutions per locus since the time of separation of the two species considered<sup>10</sup>.  $H$ , defined as  $(1/m) \sum_{j=1}^m H_j$ , is the arithmetic mean of  $H_j$  over the  $m$  loci at which  $X$  and  $Y$  are compared.

For each pair of species in a survey we thus have the values of  $H_j$  and  $I_j$  for each locus comparison, as well as  $H$ , the average

heterozygosity for all loci in the species pair, and  $I$ , the measure of genetic identity for this pair calculated over all loci.

For further analysis, each locus comparison was regarded as an entity for which values of these four variables  $H_i$ ,  $I_i$ ,  $I$  and  $H$  had been measured: all the locus comparisons for a given species pair were assigned the value of  $I$  and the value of  $H$  calculated for that species pair. The data for all surveys were then pooled and subsequently broken down into subsamples of locus comparisons for each of the different proteins scored in the surveys. Discarding some proteins because of limited data, the following quantities were calculated for each of the remaining 31 proteins, averaging over all locus comparisons contributing data for each protein: (1)  $\bar{I}_i$ , the average value of  $I_i$ ; (2)  $\bar{I}$ , the average value of  $I$ ; (3)  $\bar{H}_i$ , the average value of  $H_i$ ; (4)  $\bar{H}$ , the average value of  $H$ . Thus,  $\bar{I}_i$  and  $\bar{H}_i$  are the average locus identity and locus heterozygosity for the protein, while  $\bar{I}$  and  $\bar{H}$  are the average species identity and species heterozygosity for those species contributing data for the protein. To facilitate discussion of the results, the identity values were transformed to distance values where  $\bar{D}_i = -\log_e(\bar{I}_i)$  and  $\bar{D} = -\log_e(\bar{I})$ . The 31 proteins studied, the number of locus comparisons per protein and the number of surveys which contributed data for each protein are given in Fig. 1 legend.

The main objective in analysis was to examine the co-variation of protein heterozygosity ( $\bar{H}_i$ ) and protein genetic distance ( $\bar{D}_i$ ). A problem of interpretation could arise if a difference in  $\bar{D}_i$  between two proteins resulted because the species pairs contributing locus comparisons for the first protein were on average more distantly related and had higher  $\bar{D}$  values than the species pairs contributing data for the second protein. Two methods of analysis were used to control for the variation in  $\bar{D}_i$  caused by variation in  $\bar{D}$  between proteins. First, variation in  $\bar{D}_i$  as the dependent variable was analysed by multiple linear



**Fig. 1** Genetic distance ( $\bar{D}_i^*$ ) plotted against heterozygosity ( $\bar{H}_i$ ) for 31 proteins. The proteins are given below. The first number in parentheses after each protein is the number of locus comparisons for the protein, the second is the number of surveys contributing data for the protein, the third is the percentage of locus comparisons where  $\bar{H}_i = 0$ . The solid line is the regression of  $\bar{D}_i^*$  on  $\bar{H}_i$ . 1, alcohol dehydrogenase (137, 21, 36); 2, malate dehydrogenase (763, 53, 70); 3,  $\alpha$ -glycerophosphate dehydrogenase (338, 46, 56); 4, isocitrate dehydrogenase (422, 42, 46); 5, sorbitol dehydrogenase (112, 21, 51); 6, 6-phosphogluconate dehydrogenase (349, 45, 18); 7, lactate dehydrogenase (889, 55, 69); 8, malic enzyme (152, 20, 87); 9, glucose-6-phosphate dehydrogenase (39, 11, 80); 10, superoxide dismutase (291, 43, 72); 11, aspartate aminotransferase (583, 47, 48); 12, phosphoglucomutase (560, 50, 38); 13, 'esterases', excluding isozymes specific for umbelliferol esters (877, 43, 28); 14, glucosephosphate isomerase (375, 43, 29); 15, xanthine dehydrogenase (147, 17, 43); 16, glutamate dehydrogenase (82, 13, 90); 17, peptidase, (excluding leucine aminopeptidase) (300, 28, 33); 18, fumarase (104, 17, 64); 19, leucine aminopeptidase (183, 20, 44); 20, transferrin (147, 20, 10); 21, unspecified non-enzymatic proteins (1162, 42, 87); 22, haemoglobin ( $\alpha + \beta$ ) (176, 22, 72); 23, albumin (158, 29, 45); 24, adenylate kinase (68, 9, 90); 25, creatine kinase (126, 9, 68); 26, adenosine deaminase (30, 8, 20); 27, mannose phosphate isomerase (61, 8, 15); 28, acid phosphatase (33, 8, 67); 29, 'nothing dehydrogenase' (33, 5, 18); 30, alkaline phosphatase (57, 4, 98); 31, octanol dehydrogenase (50, 7, 54).

regression, with  $\bar{H}_i$ ,  $\bar{D}$  and  $\bar{H}$  as independent variables; this permitted a control for variation in  $\bar{H}$  as well as  $\bar{D}$ . Second, a new variable,  $\bar{D}_i^* = \bar{D}_i \times \bar{D}/\bar{D}$ , was calculated for each protein, where  $\bar{D}$  is the grand mean of  $\bar{D}$  over all proteins. The term  $\bar{D}/\bar{D}$  ensures that if the species pairs contributing data for a protein give a high or low  $\bar{D}$  value compared with the average ( $\bar{D}$ ), then  $\bar{D}_i^*$  will be adjusted to be, respectively, lower or higher than  $\bar{D}_i$ . Variation in  $\bar{D}_i^*$  was analysed by simple linear regression on  $\bar{H}_i$ .

Results of both methods of analysis are given in Table 1. Protein heterozygosity ( $\bar{H}_i$ ) accounts for about 70% of the variation in protein genetic distance ( $\bar{D}_i$  or  $\bar{D}_i^*$ ) after controlling for  $\bar{D}$ .  $\bar{D}$  accounts for a significant proportion of the variation in  $\bar{D}_i$ ;  $\bar{H}$  does not.  $\bar{D}_i^*$  is plotted against  $\bar{H}_i$  in Fig. 1. It can be seen that proteins having high heterozygosity show greater genetic divergence than proteins of low heterozygosity.

In interpreting this association, the possibility of a statistical artefact should be considered. This arises because the sampling error of genetic distance can depend not only on the number of individual animals sampled per species but also on heterozygosity, the magnitude of the sampling error increasing with increasing heterozygosity. This can lead to spurious associations between the two variables<sup>5</sup>. The analysis was, however, repeated with the possibility of a serious artefact excluded. This was done by using the original values of  $\bar{H}_i$  as characteristic heterozygosity measures of the proteins but recalculating  $\bar{D}_i$ ,  $\bar{D}$  and  $\bar{D}_i^*$  using only those locus comparisons where locus heterozygosity ( $H_i$ ) equals zero, that is, where both species were monomorphic for the locus considered. In this situation locus identity ( $I_i$ ) can only be 0 or 1 and will hardly be affected by sampling error. The  $\bar{I}_i$  value for each protein will be expressed as the simple proportion: number of cases where  $I_i = 1$  divided by number of cases where  $I_i = 0$  or 1. The results, also given in Table 1, confirm those of the first analyses, except that  $\bar{H}_i$  accounts for a smaller proportion of the variation in protein genetic distance. In addition, they suggest that contemporary differences in genetic distance between protein loci now observed as monomorphic within species have arisen because some proteins have been inherently more variable in the past than others. This, together with the observation that a large proportion of all loci are monomorphic (see Fig. 1 legend), suggests that the level of heterozygosity of individual loci varies considerably over time and that the contemporary level of heterozygosity of a locus may give a less accurate indication of future or past heterozygosity than knowledge of average heterozygosity ( $\bar{H}_i$ ) for the protein.

Published estimates of average amino acid substitution rates for 6 of the 31 proteins studied<sup>11-13</sup> are given together with  $\bar{D}_i^*$  and  $\bar{H}_i$  values in Table 2. The high and significant correlations between relative electrophoretic divergence rate, relative amino acid substitution rate and heterozygosity provide strong support for theories which relate amino acid substitution rate to

**Table 1** Results of regression of protein genetic distance ( $\bar{D}_i$  or  $\bar{D}_i^*$ ) on protein heterozygosity ( $\bar{H}_i$ ),  $\bar{D}$  and  $\bar{H}$  for 31 proteins

|                                                                               | All loci   | Monomorphic loci |
|-------------------------------------------------------------------------------|------------|------------------|
| Simple correlation of $\bar{D}_i$ with $\bar{H}_i$                            | 0.759*     | 0.576*           |
| Partial correlation of $\bar{D}_i$ with $\bar{H}_i$ controlling for $\bar{D}$ | 0.834*     | 0.621*           |
| Proportion of variation in $\bar{D}_i$ explained by $\bar{H}_i$               | 0.576*     | 0.332*           |
| Proportion of variation in $\bar{D}_i$ explained by $\bar{D}$                 | 0.198*     | 0.378*           |
| Proportion of variation in $\bar{D}_i$ explained by $\bar{H}$                 | 0.025 (NS) | 0.000 (NS)       |
| Correlation of $\bar{D}_i^*$ with $\bar{H}_i$                                 | 0.828*     | 0.686*           |
| Regression coefficient of $\bar{D}_i^*$ on $\bar{H}_i$                        | 3.240*     | 2.719*           |
| Constant for regression of $\bar{D}_i^*$ on $\bar{H}_i$                       | 0.163*     | 0.157†           |

The first column of values is for the analysis using all locus comparisons. The second column is for the analysis using only those loci monomorphic in both species, in the calculation of  $\bar{D}_i^*$ . The partial correlation has 28 degrees of freedom; other correlations have 29 degrees of freedom.

\*  $P < 0.001$ ; †  $P < 0.01$ ; NS, not significant.



**Table 2** Unit evolutionary period (UEP), protein genetic distance ( $\bar{D}_i^*$ ), protein heterozygosity ( $\bar{H}_i$ ) and relative divergence rates for six proteins

| Protein                                                                           | UEP  | Amino acid relative rate | $\bar{D}_i^*$ | $\bar{D}_i^*$ relative rate | $\bar{H}_i$ |
|-----------------------------------------------------------------------------------|------|--------------------------|---------------|-----------------------------|-------------|
| Transferrin                                                                       | 2    | 1.00                     | 1.04          | 1.00                        | 0.203       |
| Albumin                                                                           | 3    | 0.47                     | 0.65          | 0.63                        | 0.121       |
| Haemoglobin ( $\alpha + \beta$ )                                                  | 3.5  | 0.11                     | 0.46          | 0.44                        | 0.073       |
| Aspartate aminotransferase                                                        | 6–12 | 0.13                     | 0.31          | 0.30                        | 0.058       |
| Lactate dehydrogenase                                                             | 16   | 0.05                     | 0.29          | 0.28                        | 0.021       |
| Glutamate dehydrogenase                                                           | 55   | 0.03                     | 0.004         | 0.004                       | 0.015       |
| Correlation of amino acid relative rate with $\bar{D}_i^*$ relative rate = 0.934* |      |                          |               |                             |             |
| Correlation of amino acid relative rate with $\bar{H}_i$ = 0.972*                 |      |                          |               |                             |             |
| Correlation of $\bar{D}_i^*$ relative rate with $\bar{H}_i$ = 0.971*              |      |                          |               |                             |             |

The unit evolutionary period is the average time, in millions of years, for a 1% difference in amino acid sequence to arise between two lineages. UEP values for albumin, haemoglobin, lactate dehydrogenase and glutamate dehydrogenase were obtained from ref. 12; the values for transferrin and aspartate aminotransferase were estimated from information given in refs 11 and 13, respectively. A value of 9 for the UEP of aspartate aminotransferase was assumed for computation of the product moment correlation coefficients. Amino acid relative rate = (UEP (transferrin) × subunit molecular weight (protein X)) / (UEP (protein X) × subunit molecular weight (transferrin));  $\bar{D}_i^*$  relative rate =  $\bar{D}_i^*$  (protein X) /  $\bar{D}_i^*$  (transferrin). Subunit molecular weight is introduced into the equation for amino acid relative rate to correct for the fact that, unlike  $\bar{D}_i^*$ , UEP is based on a percentage difference. The correlation coefficients have four degrees of freedom.

\*  $P < 0.01$ .

heterozygosity. Furthermore, the high correlation between the two different measures of evolutionary rate suggests that Fig. 1 provides a reasonable indication of amino acid substitution rate for a number of proteins not yet fully investigated by amino acid sequencing.

The finding that variation in  $\bar{D}_i$  also accounts for a significant proportion (~20%) of the variation in  $\bar{D}_i$  among proteins can be attributed to the observation that different selections of proteins are used in different electrophoretic surveys. The implications of this for phylogenetic studies have been considered previously. For example, a bimodality of electrophoretically measured rates of evolution at protein loci has been proposed, with one set of proteins including albumin, transferrin and nonspecific esterases evolving 10 times faster than other proteins<sup>11</sup>. It was suggested that setting of any electrophoretic clock should thus take account of the sample of proteins used for comparison. Our results also indicate that this is desirable, but although albumin, transferrin and esterases have high genetic distance values, no obvious bimodality is suggested by Fig. 1 and there is therefore no justification for a correction based on the simple expedient of dividing the proteins into two groups. Each protein should be calibrated separately.

The results may also throw light on the problem of the driving force of molecular evolution. In neutral theory, genetic distance ( $D$ ) can be approximated by a function of mutation rate ( $u$ ) and the period of time since a pair of species became isolated ( $t$ ); heterozygosity ( $H$ ) can be approximated by a function of  $u$  and effective population size ( $N_e$ ). This is true both for the infinite allele model<sup>10,14</sup> and the stepwise mutation model<sup>15,16</sup>. By substitution for  $u$ ,  $D$  can be expressed in both models as a function of  $H$ ,  $t$  and  $N_e$ . For example, in the infinite allele model the equations  $D = 2ut$  (ref. 10) and  $H = 1 - 1/(4N_e u + 1)$  (ref. 14) lead to  $D = Ht/2N_e(1 - H)$ , or with respect to the present study,  $\bar{D}_i^* = \bar{H}_i t / 2N_e(1 - \bar{H}_i)$ . In both models, for specific values of  $N_e$  and  $t$ , the relationship between  $\bar{D}_i^*$  and  $\bar{H}_i$  is approximately linear in the heterozygosity range of Fig. 1, and gives  $\bar{D}_i^* = 0$  when  $\bar{H}_i = 0$ . If different values of  $t$  or  $N_e$  are chosen, only the slope is altered. In the present analysis, each protein has data sampled from the same compilation of allozyme surveys and thus the averages of  $t$  and  $N_e$  should be similar for all proteins studied. Different proteins do not, however, have data contributed by exactly the same set of surveys, and use of  $\bar{D}_i^*$  rather than  $\bar{D}_i$  provides an additional control for  $t$ . If the observed results do agree with neutral theory, it should be possible to draw just one theoretical line giving a good fit to the scatter of points in Fig. 1. This cannot be done, regardless

of the values given to  $t$  and  $N_e$ , because the observed linear relationship will, if extended back, cut the ordinate: any theoretical line derived from neutral theory is approximately linear but must pass through the origin. Perhaps the best strategy to explain the results with neutral theory is to choose values of  $t$  and  $N_e$  such that the theoretical relationship passes through the origin and somewhere close to  $\bar{D}_i^* = 0.8$  and  $\bar{H}_i = 0.2$ , thus accounting for the high distance values for the most highly heterozygous proteins; under neutral theory such proteins are thought to evolve rapidly as a result of low functional and selective constraints<sup>12</sup>. If this is done, many low heterozygosity proteins have higher genetic distance than predicted from neutral theory. The excess genetic distance can, however, be explained as the result of positive directional selection for advantageous mutants, because for a given contribution to contemporary heterozygosity the substitution rate can be much higher for advantageous than neutral mutants<sup>17</sup>. With this model, many proteins in the heterozygosity range 0.00–0.04 will have accumulated more selective than neutral substitutions. Note, however, that even if only a relatively small proportion of substitutions occur as a result of fixation of neutral mutants, the vast majority of contemporary polymorphisms will be neutral<sup>17,18</sup>. The observation of significant heterogeneity in genetic distance in this range (for example, alkaline phosphatase and adenylate kinase have  $\bar{D}_i^*$  values of 0.014 and 0.343 with standard errors<sup>19</sup> of 0.016 and 0.078, respectively) also indicates selection of some kind. The modified neutral model in which a proportion of variation is maintained by mutation–selection balance<sup>20</sup> could explain the relatively high heterozygosity and low genetic distance values of some proteins, for example, glutamate dehydrogenase and alkaline phosphatase.

It thus appears that the results can be explained by a combination of neutral and selective substitutions, but not by neutral substitutions alone. However, the positive slope of Fig. 1 is also consistent with a purely selectionist interpretation where, for example, higher levels of balanced polymorphisms and increased substitution rates could occur together in those proteins subject to less stringent functional and structural constraints. However, even a negative slope or a random scatter of points in Fig. 1 could have been accounted for by assuming appropriate differences among proteins in the propensity to maintain polymorphisms, and in the probability that in a given time such maintained polymorphisms are converted into transient polymorphisms. Neutral theory is therefore much more powerful in providing clear predictions which can be tested by these results.

We thank R. A. Galleguillos, P. Sattler and G. Wallis for providing unpublished data and M. Viner for advice on methods of computer analysis.

Received 16 November 1981; accepted 17 May 1982.

1. Sokal, R. R. in *Ecological Genetics: The Interface* (ed. Brussard, P. F.) (Springer, New York, 1978).
2. Riska, B. *Evolution* **33**, 1001–1004 (1979).
3. Johnson, M. S. & Mickevich, M. F. *Evolution* **31**, 642–648 (1977).
4. Pierce, B. A. & Mitton, J. B. *Syst. Zool.* **28**, 63–70 (1979).
5. Skibinski, D. O. F. & Ward, R. D. *Genet. Res.* **38**, 71–92 (1981).
6. Selander, R. K. in *Molecular Evolution* (ed. Ayala, F. J.) (Sinauer Associates, Sunderland, 1976).
7. Fuerst, P. A., Chakraborty, R. & Nei, M. *Genetics* **86**, 455–483 (1977).
8. Chakraborty, R., Fuerst, P. A. & Nei, M. *Genetics* **88**, 367–390 (1978).
9. Ward, R. D. *Biochem. Genet.* **16**, 799–810 (1978).
10. Nei, M. *Am. Nat.* **106**, 283–292 (1972).
11. Sarich, V. M. *Nature* **265**, 24–28 (1977).
12. Wilson, A. C., Carlson, S. S. & White, T. J. *A. Rev. Biochem.* **46**, 573–639 (1977).
13. Sonderegger, P. & Christen, P. *Nature* **275**, 157–159 (1978).
14. Kimura, M. & Crow, J. F. *Genetics* **49**, 725–738 (1964).
15. Chakraborty, R. & Nei, M. *Evolution* **31**, 347–356 (1977).
16. Ohta, T. & Kimura, M. *Genet. Res.* **22**, 201–204 (1973).
17. Maruyama, T. *J. molec. Evol.* **1**, 368–370 (1972).
18. Nei, M. *Molecular Population Genetics and Evolution* (North-Holland, New York, 1975).
19. Nei, M. *Am. Nat.* **105**, 385–398 (1971).
20. Kimura, M. *Proc. nat. Acad. Sci. U.S.A.* **76**, 3440–3444 (1979).

## MATTERS ARISING

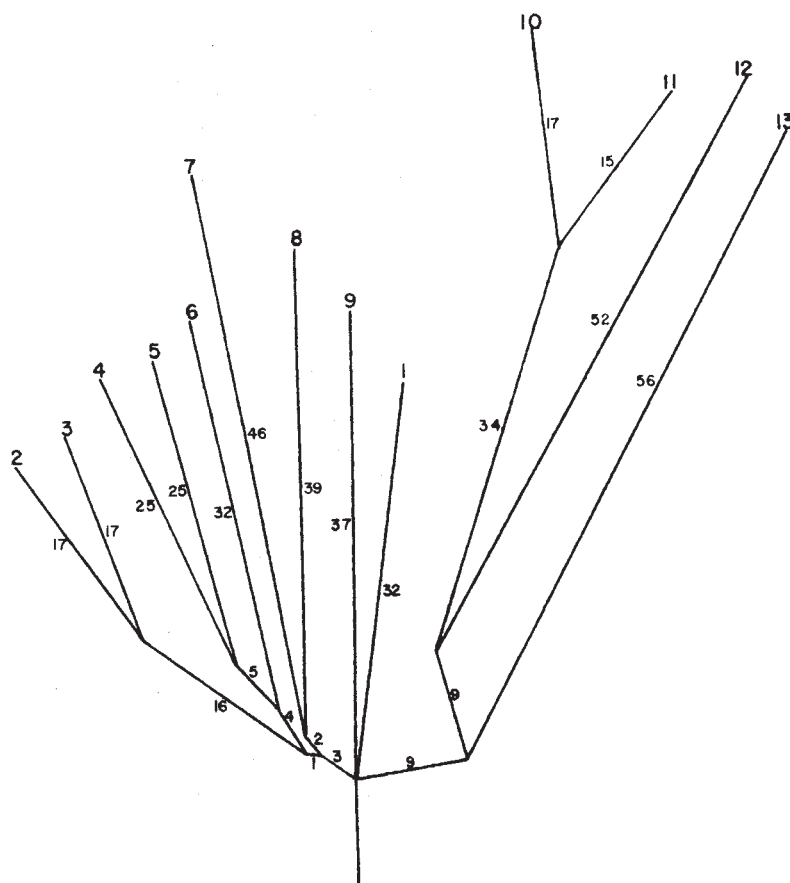
Phylogenies in molecular evolution: *Prochloron*

SIMILARITY is a mirror for phylogeny only when evolutionary rates are constant, and parallelism, convergence and reversals can be ignored. This problem is particularly acute when phylogenies are inferred from sequences of proteins or nucleic acids, where there is ordinarily no auxiliary information available (or at least used) to determine the direction of evolution at homologous sites in the molecule. Seewaldt and Stackebrandt<sup>1</sup> have constructed a phylogeny of proalgae<sup>2</sup> and chloroplasts from the similarity of fragments of 16S rRNA and conclude that *Prochloron* is well within the radiation of blue-green algae. I have constructed an alternative phylogeny by iteration from their data and conclude that *Prochloron*

**Table 1** Distances between taxa, from Fig. 1 (upper right) and from Seewaldt and Stackebrandt<sup>1</sup> (lower left)

| Taxon | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 |
|-------|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1     |    | 54 | 54 | 55 | 55 | 57 | 65 | 61 | 57 | 69 | 68 | 73 | 73 |
| 2     | 51 |    | 31 | 52 | 52 | 54 | 63 | 59 | 58 | 70 | 69 | 73 | 73 |
| 3     | 51 | 30 |    | 52 | 52 | 54 | 63 | 59 | 58 | 70 | 69 | 73 | 73 |
| 4     | 53 | 53 | 53 |    | 44 | 52 | 64 | 60 | 59 | 70 | 69 | 74 | 74 |
| 5     | 53 | 53 | 53 | 42 |    | 52 | 64 | 60 | 59 | 70 | 69 | 74 | 74 |
| 6     | 53 | 53 | 53 | 50 | 50 |    | 66 | 61 | 61 | 72 | 71 | 75 | 75 |
| 7     | 65 | 65 | 65 | 65 | 65 | 65 |    | 67 | 68 | 77 | 76 | 80 | 79 |
| 8     | 60 | 60 | 60 | 60 | 60 | 60 | 65 |    | 63 | 74 | 73 | 77 | 77 |
| 9     | 62 | 62 | 62 | 62 | 62 | 62 | 65 | 62 |    | 71 | 71 | 75 | 75 |
| 10    | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 |    | 29 | 74 | 78 |
| 11    | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 30 |    | 73 | 78 |
| 12    | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 |    | 81 |
| 13    | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 |    |

The distance between any two taxa is the sum of the corrected lengths of the segments of the estimated phylogenetic tree connecting the taxa. In order to combine any pair of distances along the segments of my phylogeny, I subtracted their product from their sum while still keeping the distances in their original form, that is, decimal fractions. This procedure makes the distances as additive as the data themselves permit. Taxa are numbered according to Fig. 1 and ref. 1.



**Fig. 1** Estimated phylogeny of some proalgae and chloroplasts. The number alongside each segment represents the distance along that segment. That the chloroplast of the red alga *Porphyridium* no longer falls within the radiation of the blue-green algae studied comes from fitting the data, not as a planned constraint. Taxon numbers are Prochlorophyta: 1, *Prochloron*. Cyanophyta: 2, *Nostoc*; 3, *Fischerella*; 4, *Aphanocapsa* 6714; 5, *Agmenellum*; 6, *Aphanocapsa* 6701; 7, *Synechococcus* 6301; 8, *Synechococcus* 7502. Rhodophyta: 9, *Porphyridium*. Angiospermae: 10, *Zea*; 11, *Lemna*. Chlorophyta: 12, *Chlamydomonas*. Euglenophyta: 13, *Euglena*.

is more probably a comparatively little modified descendant of the common ancestor of blue-green algae and chloroplasts of green plants.

The basis for my phylogeny is the usual view<sup>2-6</sup>, which I justify below, that chloroplasts of green plants are descended from the Prochlorophyta, a group of proalgae for which *Prochloron* is the only known surviving member. I used this conclusion as my only constraint in deriving a phylogeny. *Prochloron* (whose main characters Raven<sup>7</sup> predicted before its discovery) contains chlorophyll *b*, like green plants but unlike blue-green and red algae; it lacks the phycobiliproteins and phycobilisomes characteristic of the latter groups; and it has paired thylakoids resembling those of green plants. Figure 1 shows my phylogeny and Table 1 gives the distances along its branches and along those of Seewaldt and Stackebrandt's phylogeny. If  $D_{Si}$  is the distance between the  $i$ th pair of species in the latter phylogeny,  $D_{Vi}$  is the respective distance in my phylogeny (see Table 1) and  $D_i$  is the distance observed [ $D_i = 100(1 - S_i)$ , where  $S_i$  is the  $i$ th binary matching coefficient in Table 1 of Seewaldt and Stackebrandt<sup>1</sup>], then

$$\sum |D_{Si} - D_i| - \sum |D_{Vi} - D_i| \\ = 157 - 141 = 16$$

Thus my phylogeny is in slightly better agreement with the data of Seewaldt and Stackebrandt than is their own phylogeny, although the difference is undoubtedly well within the limits of error of estimation.

The rRNA data are thus fully consistent with very different phylogenies. This con-



clusion can be generalized beyond the data of Seewaldt and Stackebrandt<sup>1</sup>; by itself, similarity is a dangerous criterion for phylogeny construction although it can often exclude many possibilities. In particular, the related methods of Li<sup>8</sup> and Klotz and Blanken<sup>9</sup> correct for unequal rates of evolution although they cannot find the root of the phylogeny without making an assumption on rates. Their methods are probably the best available for constructing phylogenies from molecular sequences. If we can find evidence on the polarity of change in some characters we can usually do better, and here rate of evolution does not confound the analysis at all.

Although the phylogenies are about equally plausible on the basis of the rRNA data, I consider that the one I give here gains greater plausibility from its concordance with the direction of evolution of other characters<sup>2</sup>. Chlorophyll *b* (with its associated protein) and paired thylakoids are shared derived characters of *Prochloron* and green chloroplasts while phycobiliproteins and phycobilisomes are shared derived characters of blue-green algae and red chloroplasts. A common ancestor presumably lacked each of these characters.

The 16S rRNA of *Prochloron* is indeed most similar to that of *Nostoc* and *Fischerella* (and *Agmenellum*), but this does not seem to reflect the pattern of phylogeny; rather, all four genera are less divergent from their common ancestor than are other genera. *Prochloron* thus seems to have undergone relatively little evolution in its rRNA since it diverged from the ancestors of blue-green algae and green chloroplasts. Possibly this may prove true also for other aspects of its phenotype.

I thank R. S. Alberty, J. L. King, R. Lande, V. C. Maiorana and E. Stackebrandt for comments.

LEIGH M. VAN VALEN

Biology Department (Whitman),  
University of Chicago,  
915 East 57th Street,  
Chicago, Illinois 60637, USA

## Female-biased sex ratios

COLWELL<sup>1</sup> has recently argued that the evolution of female-biased sex ratios, in situations where there is local competition between males for access to females (Hamilton's 'local mate competition'<sup>2,4</sup>), depends critically on the operation of group selection. His model assumes that, in each generation, the population consists of *n* fertilized females drawn at random from an infinite pool; mating takes place within each group, followed by dispersal of the females to produce the pool from which the next generation is formed. The selective advantage of a genotype causing a female-biased sex ratio among its progeny arises from the fact that males do not mate outside their group, and thus contribute to the gene pool only through the genes carried by the progeny of the female members of their own group<sup>1-4</sup>. Since, in this model, the individual groups have no permanent existence, the gene pool relevant to calculations of the effect of selection is necessarily that of the progeny of the dispersing females. It is hard to see why this should be called group selection, as the term is usually reserved for situations in which selection operates on the differential contributions of partially isolated populations with some degree of permanence<sup>5</sup>. (The model of Bulmer and Taylor<sup>4</sup>, in which groups persist for more than one generation, meets this criterion.)

Female-biased sex ratios could indeed evolve if there were no discrete groups, but individuals were distributed over a spatial continuum, with females dispersing every generation after mating with males from the same neighbourhood. Moreover, Hamilton's principle also applies to the division of reproductive resources between male and female function in hermaphrodites reproducing by a mixture of self-fertilization and outcrossing<sup>6</sup>, where the question of group selection does not arise. Similarly, competition between related individuals of one sex for resources other than access to members of the opposite sex can generate biased sex ratios in random-mating populations<sup>7,8</sup>. The critical factor is thus the existence of more severe competition between relatives of one sex compared with the other, rather than competition between breeding groups. On Colwell's argument, any kind of frequency-dependent selection involving interactions between neighbouring individuals would logically be classed as group selection; this seems to us to blur an important distinction concerning the level of the reproductive units on which selection acts<sup>5</sup>.

B. CHARLESWORTH  
M. A. TORO

School of Biological Sciences,  
University of Sussex,  
Brighton BN1 9QG, UK

1. Colwell, R. K. *Nature* **290**, 401-404 (1981).
2. Hamilton, W. D. *Science* **156**, 477-488 (1967).
3. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
4. Bulmer, M. G. & Taylor, P. D. *J. theor. Biol.* **86**, 83-89 (1980).
5. Maynard Smith, J. Q. *Rev. Biol.* **51**, 277-283 (1976).
6. Charlesworth, D. & Charlesworth, B. *Biol. J. Linn. Soc.* **15**, 57-74 (1981).
7. Clark, A. B. *Science*, **201**, 163-165 (1978).
8. Toro, M. A. *J. theor. Biol.* **95**, 305-311 (1982).

RECENTLY Colwell<sup>1</sup> has claimed that Hamilton's<sup>2,3</sup> model for the evolution of female-biased sex ratios in haplodiploid arthropods, although correct in its predictions, does not correctly identify the causative factor effecting these biases. Hamilton<sup>2,3</sup> has argued that female-biased sex-related investment patterns are a result of inbreeding and what he calls 'local mate competition' (LMC). Colwell has developed a model which shows that female-biased sex ratios can evolve through group selection in highly ephemeral subpopulations. For groups in which the number of foundresses is greater than one, he claims that this "... indicates that (1) LMC itself cannot select for female biased progenies, (2) inbreeding (sib mating) is not relevant to the problem, and (3) group selection (differential productivity of genetically different groups) is required for the evolution of female-biased sex ratios". These claims are significant both because Hamilton's work<sup>2,3</sup> is generally accepted as a definitive explanation of sex-related investment biases<sup>4</sup> and because it forms the basis for several important models for the evolution of social<sup>5,6</sup> and genetic<sup>6</sup> systems.

Colwell's model shows that in restricted conditions (also noted by Hamilton<sup>2</sup>), group selection can lead to female-biased sex ratios. Yet, he neither offers evidence that group selection is the only mechanism for the evolution of such biases, nor refutes Hamilton's claim that the high likelihood of a son's success in inseminating his sisters forms the basis for shifts of maternal investment away from sons in favour of daughters.

The group selection model proposed by Colwell does not account for an important feature of populations where investment biases are common: extreme sex ratios favouring females even in situations where many inseminated females are close together. The group selection model predicts that extreme sex-ratio biases will characterize only very small populations. Thus, the existence of females producing biased sex ratios at population densities much above those predicted by the group selection model not only fails to support that model<sup>8,9</sup> but also lends additional support to inbreeding within families as a cause of sex-ratio biases.

Evidence in species where female-biased sex ratios are common supports the view that inbreeding among immediate relatives is common. This includes evidence showing sib mating within the

1. Seewaldt, E. & Stackebrandt, E. *Nature* **295**, 618-620 (1982).
2. Van Valen, L. M. & Maiorana, V. C. *Nature* **287**, 248-250 (1980).
3. Withers, N. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2301-2305 (1978).
4. Giddings, T. H., Withers, N. W. & Staehelin, L. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 352-356 (1980).
5. Margulis, L. *Symbiosis in Cell Evolution* (W. H. Freeman, San Francisco, 1981).
6. Lewin, R. A. *Ann. N.Y. Acad. Sci.* **361**, 325-329 (1981).
7. Raven, P. H. *Science* **169**, 641-646 (1970).
8. Li, W. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1085-1089 (1981).
9. Klotz, L. C. & Blanken, R. L. *J. theor. Biol.* **91**, 261-272 (1981).

mother<sup>8,9</sup>, male guarding of immature sisters<sup>10</sup>, the production of weak males<sup>3</sup> and mother-son matings<sup>1</sup>.

The difficulty with Colwell's analysis emerges from his failure to consider predictions from his model in two situations. The first arises when the size of reproductive and spacial groups is not equal. Colwell defines the conditions for his model in terms of the number of foundresses per group. However, for obligate inbreeders, the number of foundresses per spacial group (as used in classical group selection models) is largely irrelevant; the size of the reproductive unit or group will determine female investment strategies. Sons of females producing different sex ratios of offspring in families with close inbreeding are not in direct competition for copulations. Without this competition, none of the losses Colwell describes will occur for mothers investing more in daughters when they are compared with other females in the same group that produce a more even sex ratio.

Inbreeding among sibs can directly affect sex ratios in another way. If a fraction of sibs commonly inbreed and the remainder outbreed, then sex ratios for the whole brood will show a greater bias towards females than would be expected if there were no inbreeding. The sex ratio for the inbred fraction will show an extreme bias favouring females, whereas that of the outbred fraction should reflect population structure (R. Colwell has independently reached this same conclusion). That is, in a large open population the sex ratio should approach unity, whereas for structured populations the model developed by Colwell will provide the best prediction. The whole-brood sex ratio will depend on (1) the proportion of females that breed with brothers, (2) the effect of a male's earlier matings on his ability to mate again, (3) the predictability of sib matings in successive generations, and (4) the degree of synchrony in achieving sexual competence among members of a brood. For groups with partial inbreeding among sibs, biased sex ratios will occur even where spacial groups have more than one foundress. Thus with evidence that inbreeding is an important factor in populations with female-biased sex-related investment ratios, group selection of the type described by Colwell is only one of several possible mechanisms for the evolution of female-biased sex ratios, and the extent of its importance in natural populations remains to be determined.

GERALD BORGIA

Department of Zoology,  
University of Maryland,  
College Park,  
Maryland 20742, USA

3. Hamilton, W. D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. & S. Blum) 167-221 (Academic, New York, 1979).
4. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
5. Hamilton, W. D. *A. Rev. ecol. Syst.* 3, 193-232 (1972).
6. Alexander, R. D. & Sherman, P. W. *Science* 196, 494-500 (1977).
7. Borgia, G. *Theor. Pop. Biol.* 17, 103-128 (1980).
8. Newstad, R. & H. Duvall, R. *Soc. Rev. no. 2 of Grain Pests (War) Committee*, London (1918).
9. Cooper, K. *Proc. natn. Acad. Sci. U.S.A.* 23, 41 (1937).
10. Cowan, D. thesis, Univ. Michigan (1978).
11. Van Emden, J. Z. *Morph. Ökol. Tier.* 23, 425 (1981).

COLWELL<sup>1</sup> has used a population-genetic argument to infer that the occurrence of a female-biased sex ratio results from a balance between individual and group selection. This result is compatible with the population homeorheostat model that I proposed previously<sup>2</sup>, but may refer to a specialized case applicable particularly to the parasitoid habit. The more general case is found among free-living invertebrates where, for example, a polygenic system or supergene controls polymorphic sex expression<sup>3</sup>. The supergene may carry a selected physiological trait as well as selectively neutral sex genes. The selecting factor or 'environmental switch' can rapidly produce the appropriate morph from low frequencies carried within the population. The supergene is inherited with little or no crossing-over at meiosis and will persist as long as the specific environmental switch remains operative. A third mechanism for female-biased sex ratios which does not have a genetic basis is known in amphipods<sup>4,5</sup>. Endoparasitic microsporidians carried in the egg cause a female bias in the progeny, suggesting that host and endoparasite have co-evolved to produce a host population of high reproductive potential of great population advantage where group selection processes are operating.

Recent work on female-biased sex ratios (see ref. 1 and refs therein) has been mainly theoretical and has involved predictions from population-genetic models. I believe that this approach may be inappropriate because an exact genetic analysis of polygenic systems or supergenes is practically impossible to achieve and is certainly not applicable where a non-genetic mechanism applies. A general theory to explain female-biased sex ratios in invertebrate populations<sup>3</sup> requires knowledge of both the genetics of sex and the ecology of the population in question. Future work on this subject might include experimental investigation of the mechanisms involved in a variety of invertebrates, field study of their population dynamics and the likely environmental switches which control polymorphic forms in geographically isolated populations.

D. J. WILDISH

Fisheries and Environmental Sciences,  
Department of Fisheries and Oceans,  
Biological Station,  
St Andrews, New Brunswick,  
Canada E0G 2X0

1. Colwell, R. K. *Nature* 290, 401-404 (1981).
2. Wildish, D. J. *Nature* 233, 54-55 (1971).
3. Wildish, D. J. in *Advances in Invertebrate Reproduction* Vol. 1 (eds Adiyodi, K. G. & Adiyodi, R. G.) 8-24 (Int. Soc. Invert. Reprod., India, 1977).
4. Bulnheim, H.-P. *Sonderdr. Helgolander wiss. Meeresunters.* 31, 1-33 (1978).
5. Ginburger-Vogel, T. & Desportes, I. *Mar. Fish. Rev.* 41, 3-7 (1979).

COLWELL REPLIES—My model<sup>1</sup> for the evolution of female-biased progenies assumes precisely the same conditions of population structure as Hamilton's<sup>2,3</sup> and Maynard Smith's<sup>4</sup> models: each group is founded by  $n$  mated females ( $n > 1$ ); mating is random within each group among the progeny of its founders; there is no density-dependent mortality in groups; and mated daughters of the founders disperse at random to found new groups. In these conditions, all models<sup>1,2,4</sup> agree (precisely) that the sex ratio of progenies should evolve towards an optimal female bias that depends on  $n$ .

The point of my paper was to show how selection operates in this evolutionary process. Following Hamilton's<sup>2</sup> verbal interpretation of his mathematical model, it became widely believed that, in the conditions of the model, a female that produces a daughter-biased progeny (a 'hamiltonian' female<sup>1</sup>) has a fitness advantage over a female that produces an unbiased progeny (a 'fisherian' female<sup>1</sup>), because the sons of the hamiltonian female compete less among themselves for mates than do the sons of the fisherian female. Charlesworth and Toro and Borgia support this view, in the face of what I believe to be its disproof<sup>1</sup>. Perhaps I can restate the case more clearly.

I showed<sup>1</sup> that in any group founded by a combination of hamiltonian and fisherian females (the simplest case is one of each), the fitness of each hamiltonian founder is invariably less than that of each fisherian founder. Fisher's principle operates with equal force in random-mating groups of any size. Indeed, if each group were founded by precisely  $H$  hamiltonian females and  $F$  fisherian females ( $H \geq 1, F \geq 1$ ), the hamiltonian trait could neither increase when rare nor be maintained when common. The trait would be eliminated in all cases, in spite of the supposed advantage of reduced local mate competition (LMC) accruing to hamiltonian founders (an 'advantage' unchanged by the assumption of invariant group composition).

What, then, accounts for the success of the hamiltonian trait in the models<sup>1,2,4</sup>? If  $H$  and  $F$  were indeed not variable among groups, each group would export the same number of dispersing female progeny. But random group founding (a condition of the models) guarantees variation among groups in the proportion of founders that are hamiltonian. Groups founded, by chance, by a higher proportion of hamiltonian females produce a greater number

1. Colwell, R. K. *Nature* 290, 401-404 (1981).  
2. Hamilton, W. D. *Science* 156, 477-488 (1967).



of dispersing daughters than do groups founded, by chance, by a lower proportion of hamiltonian females. Even though each hamiltonian female loses out locally to each fisherian female within every mixed group, the differences among groups in the production of dispersers favour the hamiltonian trait. The among-group advantage of female-biased progenies exceeds the disadvantage of this trait to individuals within groups, and the trait increases in frequency in the global population<sup>1</sup>. The optimal female bias is more extreme when groups are smaller not because LMC is stronger (though it is), nor because inbreeding is greater (though it is) in smaller groups, but because smaller groups vary more in initial composition, and thus in the number of dispersers they export<sup>1,3,5</sup>.

In the conditions of the models, it is impossible for female bias to evolve without the differential productivity of groups. This fact can be demonstrated by eliminating differential productivity, either by imposing a within-group carrying capacity<sup>5</sup> or by suppressing variation in group composition, as argued above. By any reasonable definition, differential productivity of groups is the driving force of group selection—it is to groups what classical fitness is to individuals. I agree with Charlesworth and Toro that the critical role of group selection in the evolution of female-biased sex ratios is more apparent when groups are isolated from one another for more than a single generation of within-group mating. (In fact, D. S. Wilson and I<sup>5</sup> produced such a model before I developed the one-generation model that is the focus of the present controversy<sup>1</sup>, and Bulmer and Taylor<sup>6</sup> independently derived a related multigeneration model.) But the selective forces acting in every generation of local mating—including the first—are identical, so there is no logical justification for drawing the line between one and two generations. Charlesworth and Toro also point out (apparently as a criticism) that the same evolutionary process<sup>1</sup> would operate when 'groups' have poorly defined boundaries. No doubt it does. The discrete 'group' of group selection is a conceptual

and mathematical convenience that should disturb no one used to dealing with theories of 'demes', 'populations', or 'communities'. I am persuaded that group selection should be defined not by the characteristics of groups, but by the characteristics of the process<sup>5,7-9</sup>.

Borgia's belief (following Maynard Smith<sup>4</sup>) that inbreeding (sib mating) itself selects for female-biased sex ratios is mistaken, in the conditions of all published models<sup>1-7</sup>. The argument is the same as for LMC: hamiltonian females indeed reduce the probability of redundant sib matings among their progeny, but nonetheless lose out to fisherian females in every mixed group. Inbreeding is greater in a group-structured deme than in a panmictic deme, but this is only an incidental fact, as far as sex ratios are concerned. I made no claim that "group selection is the only mechanism" for the evolution of female-biased sex ratios (Borgia), only that group selection is required in the conditions of existing models. For obligate sib mating ( $n = 1$ ), or for some of the mechanisms discussed by Wildish, group selection is clearly not required. Borgia complains that I have neglected cases in which the size of 'spatial groups' is greater than the size of 'reproductive groups'. If sib mating is not obligate, but is commoner than expected for random mating, given  $n$ , then the effective group size  $n'$  satisfies  $1 < n' < n$ , and the evolutionarily stable proportion males is correspondingly lower. (Variable mating success among males has the same effect<sup>5</sup>.) There is no published model for the case, suggested by Borgia, of partially obligate sib mating, although I have derived one (unpublished). In it there are three selective forces at work: individual selection favouring female bias, based on forced sib mating (selection against LMC); individual selection against female bias among random-mating progeny (Fisher's principle); and selection between groups for productivity, favouring female bias. The magnitude of each force and the resultant balance depend on the biology of the organism. Only the group selection term, however, depends on group size ( $n$ ). Therefore any data showing a positive correlation between number of founders and proportion males<sup>5</sup> are evidence for group selection, whatever the proportion of obligate sib matings (assuming differential mortality is uncorrelated with group size).

Wildish adds interesting detail to the complex biology of female-biased populations, but I disagree with some of his interpretations. For the purpose of generality, the model I presented<sup>1</sup> relied on no particular mode of inheritance: it was explicitly based on the fitness of phenotypes, not genes or genotypes. The mode of transmission of the sex-ratio trait is unspecified because it is essentially irrelevant to the phenomenon the model demonstrates, contrary to Wildish's inter-

pretation. I would also dispute the notion that any gene affecting progeny sex ratio (more generally, investment ratio) could possibly be "selectively neutral" (Wildish), except in a very large, panmictic population with a stable, unbiased sex ratio<sup>10</sup>. For example, in the case of a panmictic population with a sex ratio biased by linkage of a 'sex-ratio' gene with a 'survival' gene, a mutant that retains the latter while producing an unbiased (fisherian) sex ratio among its progeny will have a great selective advantage.

ROBERT K. COLWELL

Department of Zoology,  
University of California,  
Berkeley, California 94720, USA

1. Colwell, R. K. *Nature* **290**, 401–404 (1981).
2. Hamilton, W. D. *Science* **156**, 477–488 (1967).
3. Hamilton, W. D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 167–221 (Academic, New York, 1979).
4. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
5. Wilson, D. S. & Colwell, R. K. *Evolution* **35**, 882–897 (1981).
6. Bulmer, M. G. & Taylor, P. D. *J. theor. Biol.* **86**, 83–89 (1980).
7. Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, in the press).
8. Price, G. A. *Ann. hum. Genet.* **35**, 485–490 (1972).
9. Wade, M. J. *Science* **210**, 665–667 (1980).
10. Shaw, R. F. & Mohler, J. D. *Am. Nat.* **87**, 337–342 (1953).

## Errata

In the letter 'Male and Female DNAs can be discriminated using retroviral probes' by S. J. Phillips *et al.*, *Nature* **297**, 241–243 (1982), the name of the fourth author is incorrectly spelt and should read 'Eva M. Eicher'. On page 242 the ninth line of text should read 'The number of virus-related sequences in male DNA'.

In *Nature* of 10–16 June 1982, credit should have been given to Mr Gene Moore who took the photograph used on the cover. The cover is also used on the reprints of the review article 'Radar research on thunderstorms and lightning' by W. D. Rust & R. J. Doviak, *Nature* **297**, 461–468 (1982).

## Corrigendum

In the letter 'Electrophysiology of mammalian thalamic neurones *in vitro*' by R. Llinas & H. Jahnsen, *Nature* **297**, 406–408 (1982), line 4 in the right-hand column on page 407 should read 'suggesting the presence of early K conductance<sup>7</sup>, not Na conductance'.

## Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

## BOOK REVIEWS

## Dicing with reality

Leslie Orgel

THE authors of this ambitious book, a translation from the German original *Das Spiel*, first published in 1975, have tried to make accessible to the general reader a group of related ideas which elsewhere are almost always treated with considerable mathematical sophistication. They touch on a very wide range of topics, but the core of the book is concerned with the development of the size and composition of populations in time. However, this is not a book about population genetics in any conventional sense; the populations may be comprised of anything from neutrons to humans. Once the laws or rules of production (birth), transition (mutation) and decay (death) are given, the "game" is defined — the size or shape of the pieces does not affect the outcome.

The major novelty of the treatment is the replacement of probability theory, with its theorems and derivations, by board games that exemplify the most important general principles. The idea as such is not new, but the consistent application of this idea and the complete absence of formal mathematical arguments make this a unique book. I found it a lot of fun to read and expect that many others will feel the same way.

But does the approach provide new insights or make old ones more intelligible? I think that the book will prove very useful for that large section of the scientific community which tends to quit when the mathematical going gets rough. Of course, if one tries to cover such a wide area in a book of modest length most of the detail must be excluded. In this context, board games are ideal, since the rules can rarely be detailed enough to model interesting situations in depth, but can often make transparent important general principles, for example those controlling the stability or instability of population size and composition.

An example may help to show how a board game can illustrate a general result. Suppose we have a square board, say a chess board, and cover it at random with 32 black and 32 white pieces. Then we roll a pair of eight-faced dice to choose a square at random. In variant 1 of the game, the piece on the chosen square is removed and replaced by a piece of the opposite colour. In variant 2, a piece of the opposite colour (if one is present) is removed from any position on the board other than the chosen square and replaced by one of the same colour as that on the chosen square. Variant 1 is stable in the long term, in the

*Laws of the Game: How the Principles of Nature Govern Chance.* By Manfred Eigen and Ruthild Winkler. Translated by Robert and Rita Kimber. Pp.350. US ISBN 0-394-41806-9; UK ISBN 0-7139-1484-X. (Knopf/Allen Lane: 1982.) \$19.95, £14.95.

sense that even after a very long time a roughly equal number of black and white pieces will almost surely be left on the board. In the case of variant 2, we may confidently expect that sooner or later all the squares will be occupied by pieces of the same colour, but we cannot predict which colour. This is a very familiar result, but it is instructive to watch the development of the distribution of pieces on the board for a few "games".

I suspect that the mathematically inclined will still prefer the standard texts on population genetics, branching chain reactions, games theory and so on. For others, the present approach may provide a useful view of the wood, before they plunge into the trees.

While the core of this book is concerned with populations and growth, the authors

make brief forays into every corner of the scientific and general cultural scene. As might be expected, some of these appeal to the present reviewer more than others. I thought the section on entropy outstandingly good, and many of the discussions of molecular evolution illuminating, but I couldn't make out a strong logical connection between the main content of the book and some of the later sections, for example, the avuncular advice offered in the section entitled "Limits".

Many will profit from this lively book. However, a little parental guidance on the relation between simple theories and complex reality may be advisable for novices to the game. One final thought — many of those who read this book will have access to a small computer. Advice on setting up these games for a computer, and perhaps some additional games especially designed with a computer in mind, would be a welcome addition to a second edition. □

*Leslie Orgel is a Senior Fellow at the Salk Institute, San Diego, California.*

## Biogeography and natural experiments

C. D. Pigott

*The Changing Climate: Responses of the Natural Fauna and Flora.* By Michael J. Ford. Pp.190. ISBN 0-04-574017-8. (George Allen & Unwin: 1982.) £13.95, \$27.50.

ONE of the generally accepted principles of biogeography is that geographical distributions of plants and animals are primarily controlled by climate. It is therefore a fundamental weakness of the subject that so important a premise cannot readily be verified by experiment. Climate cannot easily be changed and the component factors of radiation, temperature, humidity and rainfall are physically interrelated so that their separate influences are only identified with great difficulty. Neither can whole ecosystems be moved from one climatic region to another, and experiments which consist of transferring individual organisms are so far removed from reality that the results must be treated with great caution.

It is in this context that climatic change assumes particular interest because it provides what are essentially natural

experiments. As with all experiments both the change imposed and the response of the organisms must be accurately measured. The value of Michael Ford's book is that it brings together for the first time a large number of scattered observations on the response of organisms to natural variations in climate, and that it also provides an excellent summary of the climatic changes which have occurred during the period for which meteorological measurements are available (Chapter 3). In the following chapter evidence for changes before this period is also presented, but it is important to recognize that this largely, though not entirely, consists of fossil evidence for change in the distribution of organisms. There is, therefore, an inevitable circularity of argument.

The author avoids this danger by restricting his discussion in the rest of the book to responses of plants and animals to climatic changes since about 1900. Although long-term variations during this period are small — for example a rise of mean annual temperature in the Northern Hemisphere of about 0.5°C from 1890 to



1940, followed by a fall of about  $0.3^{\circ}\text{C}$  to the present day — the variations between individual years have often been very much greater. What is known of the influence of these variations, not only of temperature but also of rainfall, on metabolism of organisms, including their growth, development and reproduction, on availability of food, on dispersal and on competition is then discussed in Chapters 6–11.

The style of the book is essentially that of a review. Numerous examples are described but analysis of cause tends to be superficial. There are occasional lapses in which correlation is confused with cause, as, for instance, in the discussion of the relation of flowering of beech (*Fagus silvatica*) to temperature in June and July of the previous year. The source of this information is also omitted but must surely be the work of Professor J.D. Matthews (*Forestry* 28, 107–116; 1955).

It is no criticism of the author that so many of the examples to which he refers are little more than simple observations and lack convincing explanations. If the book stimulates biologists to make experimental studies of even a few of the phenomena it describes, it will have done a valuable service to biogeography. □

C.D. Pigott is a Professor in the Department of Biological Sciences at the University of Lancaster.

## Protein structure: beyond pretty pictures

Nicholas G. Wrigley

*Electron Microscopy of Proteins*, Vols 1 and 2. Edited by James R. Harris. Vol. 1, pp.352, ISBN 0-12-327601-2; Vol.2, pp.316, ISBN 0-12-327602-0. (Academic: 1981.) Vol. 1 £35, \$72; Vol. 2 £30, \$61.50.

PRACTITIONERS of electron microscopy lie on a spectrum of widely different philosophy. On one hand there are those to whom the technique is entirely incidental to the subject studied, and seems to command no real interest or understanding; such people adduce a mass of poorly understood and suspect EM data to support existing ideas on function or merely to describe structure. On the other hand there are those to whom the technique is all; they produce superb data but without sufficient knowledge of the subject studied to make proper use of them.

This is probably in the nature of science itself, and it is no fault of the editor of *Electron Microscopy of Proteins* that both these extremes are represented in the first two volumes of this new series. On the contrary it is very much to his credit that he has been able to include much of the middle ground also in these first 14 articles. To this extent his aim "... to review the progress achieved in the electron microscopic study of soluble, fibrous and membrane-bound proteins" has succeeded.

The series is ambitious in its sheer breadth of subject; it could be re-titled *Biological Electron Microscopy* since virtually all of biology is concerned with proteins at some level. Apparently because of the problems of extracting promised chapters from authors, there is little ordering of this breadth into sub-topics such as soluble, fibrous or membrane-bound proteins. A reader interested in fibrous proteins must buy Vol. 1 for "Intermediate Filaments", Vol. 2 for "Fibrous Proteins of Connective Tissue", and must await Vol.3 for "Tubulin and Associated Proteins" and eventually Vol.4 for "Muscle Proteins". Since the series is not about EM techniques — though some articles include substantial technical sections — such subdivision would have been desirable and is to be hoped for in any extension of the series.

Electron microscopy of biological systems can only be justified if the aim is to understand function through structure, rather than merely to catalogue structural details. On this criterion some of the studies have no place in these volumes, notwithstanding elegant microscopy. On other grounds this very elegance justifies their presence, though a more critical appraisal of their contribution in understanding function would have helped the general reader. He can thus be misled into the impression that a welter of beautiful pictures is synonymous with profound functional understanding. However the

general high standing and quality of the authorship of these chapters allows the discerning reader to decide for himself which is which. Moreover each chapter includes a compendious bibliography of its subject, which is essential for a review series of this kind. On this showing the whole series of (at least) four volumes will amount to impressive coverage of the whole field. □

Nicholas G. Wrigley is at the Medical Research Council's National Institute for Medical Research, Mill Hill, London.

## A palpable hit

P.W. Hawkes

*Image Analysis and Mathematical Morphology*. By J. Serra. Pp.610. ISBN 0-12-637240-3. (Academic: 1982.) £48.40, \$99.50.

THE appearance of this book is something of an event in the world of image analysis. For more than a decade now, scattered ideas and results have been emerging from the Ecole des Mines at Fontainebleau, suggesting that something new was in the air. A connected account was lacking, however, and even the individual publications were easily overlooked, many being internal reports and others published in a wide variety of serials. Such diversity is no accident: it reflects the broad range of possible applications of the methods but the absence of a single easily accessible text has meant that the ideas have been far too little appreciated, outside the Seine-et-Marne at least.

With the publication of this fascinating and densely filled volume, there is no longer any excuse for ignorance of the methods. Not only is a coherent pattern imposed on the work of the Fontainebleau group and the necessary mathematical background sketched in but "40 per cent of the material of this book has not been published before".

What then does Serra offer us? His text is divided into four large parts, on the mathematical tools needed, the role of partial knowledge (with chapters on digitalization and random sampling), the criteria applied in practical image analysis (covariance, size, connectivity and the study of half-tones) and finally the use of random models. All this is brought to bear on problems of image analysis in such a way that the results are in some well-defined sense independent of or insensitive to the exact choice of image (invariance under translation), its magnification, the incomplete nature of the information

### TODAY'S SCHOOL CHILDREN

will determine if nuclear  
energy is to SURVIVE

### EARTH...

a new, richly illustrated  
24-page booklet from the  
American Nuclear Society.

### EARTH...

can bring school children the true  
story of nuclear energy in today's  
environmentally-oriented society.

### EARTH...

is available in quantity, at a low  
cost, to your company for  
distribution to eighth and ninth  
grade students.

Write today for a price schedule  
and sample copy.

Public  
Communications Dept. P  
American Nuclear Society  
555 N. Kensington Avenue  
La Grange Park, IL 60525

Circle No.01 on Reader Service Card.

furnished by the images and the niceties of limits of resolution.

Mathematical morphology is a collection of methods for extracting quantitative information from images subject to these four prerequisites. The mathematical tools required will be unfamiliar to many of the potential users of the methods:

They are, amongst others, integral geometry, geometrical probability, second order random functions, graph theory, and topologies on a locally compact separable space.

Fortunately most of the methods can be grasped and applied without a profound knowledge of all these topics, and Serra takes great care to sprinkle his discussion, however abstruse and mathematical in appearance, with down-to-earth examples.

The book is addressed to potential users ("biologists, metallographers, geologists, geographers of aerial imagery"), image analysts and mathematicians working on probability theory. Hardly surprisingly, some of these will be dissatisfied, particularly as Serra is apt to forget that his readers will frequently know nothing of the subject he is discussing. Thus the chapter on the hit-or-miss transformation, which is fundamental, opens with a page and a half of discursive reflections on ways of approaching the transformation without telling us, even in qualitative terms, what it is.

Not everyone will enjoy the jocular asides ("we hope the reader will forgive us, but such is life!") but no-one can fail to appreciate the vivid examples: a section on a statistical property begins "Recall the story of the butcher who sells lark pie". Elsewhere, the prose is occasionally turgid, all too often tending to hold up the reader by the use of unwieldy jargon such as "a second *a priori* heuristic. . ." or phrases of the type "there is absolutely no condition in the principles relevant to rotation". Some of the blame for these infelicities must fall on the various people thanked for help with the English, none of whom knows how to spell cytoplasm, Feulgen and chromatin, whose use of the comma is disconcertingly capricious and whose efforts to weed out translationese are uneven.

A last complaint concerns the references, which are treated in a most cavalier fashion: thus of the 15 that connect Piaget to Roach, only two are as they should be — all the others are in some way faulty or inadequate.

Having said all this, we must be extremely grateful that the book has been written. It contains an enormous amount of information, practical as well as theoretical, most of it new or unfamiliar — it is only because it will be heavily used for some years that these shortcomings of language and presentation matter. Singularly well produced, it is essential reading (and re-reading) for anyone involved in analysing images. [1]

Peter Hawkes is Maître de Recherches at the Laboratoire d'Optique Electronique, Toulouse.

## Beginnings for historical climatology

Stephen H. Schneider

*Climate and History*. Edited by Robert I. Rotberg and Theodore K. Rabb. Pp.280. Hbk ISBN 0-691-05331-6; pbk ISBN 0-691-00787-X. (Princeton University Press: 1981.) Hbk £14.90, \$26; pbk £4.40, \$7.75. *Climate and History: Studies in Past Climates and Their Impact on Man*. Edited by T.M.L. Wigley, M.J. Ingram and G. Farmer. Pp.530. ISBN 0-521-23902-8. (Cambridge University Press: 1982.) £30, \$59.95.

NINETEEN seventy-nine was a vintage time for those climatologists and historians interested in each other's work, as in that year two major conferences were held on the topic of "climate and history". The first meeting, in Cambridge, Massachusetts, was organized by historians Theodore Rabb and Robert Rotberg, also co-editors of the *Journal of Interdisciplinary History* — the original publication vehicle for the papers and formal commentary given in Cambridge. The second, held two months later at the University of East Anglia, was organized by the Climatic Research Unit at that university. The contributions in the resulting books — 17 in the Rotberg/Rabb volume (nearly one for every participant at the Cambridge meeting) and 21 in Wigley *et al.* (one for roughly every ten of the participants at Norwich) — can be classified into several categories.

First, there are the physical and biological scientists who consider the largely technical aspects of inferring climatic information from proxy data such as those provided by tree rings, glaciers, fossil pollen remains or time-series of the oxygen isotope ratios of shafts in limestone caves. While most of these chapters are generally well-written, and although all of them certainly present a rich variety of fundamental methods, little of the information is new; almost all of it is already available elsewhere. Nonetheless, it is valuable to have such details assembled in one place.

Played off against the contributions from the natural scientists are those in which historians (or climatologist-trained surrogate historians) discuss methods by which climatic inferences may be squeezed from documentary sources such as diaries and records of grain prices, death rates, migration and so on. While due caution is usually given to warn naive readers against coming to wild conclusions on the basis of such sources, the various authors do, by-and-large, make a good case that historians can draw upon a combination of physical and biological proxy data to help them reconstruct past climates, just as palaeoclimatologists can apply — carefully, of course — historical documents to the same end.

The review articles in Wigley *et al.* on the

use of stable isotope data (by J. Gray), glaciological data (S. Porter) and pollen analyses (H. Birks) in palaeoclimatic reconstruction are succinct but comprehensive, and should be helpful to those wanting entry points to the literature and methods of these respective subdisciplines. Perhaps Birks is too sceptical about the value of multivariate, regression techniques in drawing climatic inferences from pollen counts, but his views are appropriately counterbalanced by the fairly convincing results obtained by this technique and given in Thompson Webb's article in Rotberg/Rabb. Finally, Hubert Lamb, inspirational figure behind the East Anglia meeting, provides a remarkable table which condenses into five pages virtually all the available information on principal data sources for reconstruction of past climates — it would literally take a book or two to amplify all of these entries.

So far, the "and" has been appropriately placed between the distinct subdisciplines, climate and history. But there is truly interdisciplinary thinking represented in these two volumes — even, perhaps, evidence of the emergence of a new, distinct subject. These pieces address the fundamental intellectual issue: to what extent has climate change influenced the course of history? Both books include contributions which bring us a long way from the simplistic doctrine of climatic determinism espoused some 50 years ago by Elsworth Huntington and from the ideological counter-strike of — largely Marxist — social critics who lay all blame for the social effects of climatic variations on the doorstep of inappropriate social structure.

The sharpest discussion of the relative impact of climate on history is given (in Rotberg/Rabb) by Jan DeVries, a historian. This judgement as to the importance of DeVries's chapter is supported by the fact that his article is quoted many times by the competition. For instance, in their excellent and badly needed survey article entitled "Past Climates and Their Impact on Man: A Review", which provides the best overview I have seen of the range of topics normally found in any "climate and history" study, Ingram, Farmer and Wigley echo DeVries's criticism of those workers who have assumed cause and effect between coincidental climatic extremes and documented social events,

Since they are unable to show what proportion of such misfortunes they describe are

Stephen H. Schneider is at the National Center for Atmospheric Research, Boulder, Colorado. (NCAR is sponsored by the National Science Foundation. Any opinions, findings or recommendations expressed herein are those of the author and do not necessarily reflect the views of the National Science Foundation.)



# THE FINAL EPIDEMIC:

**Physicians and Scientists  
on Nuclear War**

Edited by Ruth Adams  
and Susan Cullen

*"An inescapable lesson of contemporary medicine is that when treatment of a given disease is ineffective or where costs are insupportable, attention must be given to prevention."*  
— from the introduction.

*This book was written by distinguished physicians, lawyers, psychiatrists, scientists, and economists, and is the result of a series of symposia on the medical effects of nuclear weapons and war.*

## Available from:

*The Bulletin of the  
Atomic Scientists*  
1020 East 58th Street  
Chicago, Illinois 60637  
U.S.A.

**\$4.95 plus \$1.00 for  
postage and handling**

Only prepaid orders accepted.  
Checks must be in U.S. dollars  
and bear the name of a U.S. bank.

attributable to climate as opposed to such non-climatic factors as the dislocations and readjustments in trade and industry . . . [p.31 of Wigley *et al.*].

In his own article, DeVries argues that we should shift the emphasis from "measuring the harm done" from some short-term climatic event to the study of "the process of adaptation" of societies to environmental changes over the long term. I could not agree more.

The issue of whether technology and social organization can increase or lessen vulnerability to climatic shocks or trends has been examined in several specific case studies by a Clark University team headed by the social geographer Robert Kates. A concise summary of their findings — and thoughtful accounts of the intellectual issues — make an appropriate last chapter to the Wigley *et al.* volume. (I say this in spite of some annoyance in seeing my own views taken too far towards the climatic determinist camp; I have never attributed the apparent increase of US grain yields in the period 1958–1973 as "primarily due to an unusual streak of very favorable weather", but rather have argued in several places that this period of "favorable weather" is largely responsible for decreased interannual variability in the absolute amount of US crop yields.) Although claiming no general conclusions, the Clark group is on the right track by mining specific examples for empirical information to test hypotheses under varying social and geographical contexts. I hope that this approach will be adopted by other researchers and applied to very different case studies.

Despite the padding of a good deal of unoriginal material, these two volumes do contain enough new and thoughtful information to make them indispensable to any serious student of the multidisciplinary topic of "climate and history" — or of the emerging discipline of "historical climatology", cited in the abstract of the piece by Sharon Nicholson. But I hope that future editors of such books will ask their authors to probe more deeply into the underlying intellectual issues. First, as I have already mentioned: to what extent has climate and climatic change influenced people and institutions? And second, perhaps more importantly: what can be learned from the past which might make societies of the future less vulnerable and more adaptable to the range of climatic fluctuations which physical science tells us is possible? In this connection it is disappointing that one such proposal put forward elsewhere by Rabb, in the context of assessing the social impact of CO<sub>2</sub>, has subsequently been ideologically black-balled by US DOE administrators who took power in the Reagan administration. Despite the temporary setback to funding caused by such narrow-thinking, I believe that the study of adaptive mechanisms to deal with climatic change has such a high payoff that the emerging field of "climate

and history" may well soon become self-sustaining.

These two important volumes have not quite carried us into the territory of a new climate-history discipline; but if one is ever to emerge, I'll wager that these works will be regarded, despite their flaws, as foundation stones. □

## Roles for centrioles

M.A. Sleight

*The Centriole: A Central Enigma of Cell Biology.* By D.N. Wheatley. Pp.232. ISBN 0-444-80359-9. (Elsevier/North-Holland Biomedical: 1982.) Dfl.190, \$88.50.

TEXTBOOKS of cell biology refer to centrioles as small bodies that act as foci for mitotic spindle formation in animal cells; many proceed to indicate that centrioles may also act as ciliary basal bodies. In *The Centriole*, Dr Wheatley examines the evidence for these and other beliefs and generalizations. He concludes that although centrioles are essential for the formation of cilia or flagella, we have little understanding of any role for them in the mitotic spindle.

Experimental studies with laser-irradiated cells, and observations that the mitotic spindle of cells of higher plants, yeasts and many fungi lack centrioles, show that centrosomes are necessary for the formation and correct functioning of the spindle, but that centrioles are not essential components of centrosomes, although they may be present. One might add that ciliates lack spindle centrioles although they possess numerous ciliary "centrioles".

Many interesting questions are raised in the book — the bipolar nature of mitotic division; the origin, nine-fold symmetry and *de novo* appearance of centrioles; the nature of the centrosome; the production and role of the primary cilium present on cells of almost all tissues in the mammalian body; the concept of microtubule organizing centres; and the probable absence of DNA and likely presence of RNA in the centriole. Discussions of these and many related topics are supported by well-selected references, and the summaries at the end of many sections are especially helpful. The result is a thought-provoking and interestingly-written series of essays.

Although the abbreviated index, poor proof-reading and lack of scales on many micrographs are regrettable, the book is certainly recommended reading for specialists concerned with cell division, microtubules or cilia. □

*M.A. Sleight is Professor in and Head of the Department of Biology at the University of Southampton.*